X-ray diffraction studies of Deoxyribose Nucleic Acid

Gosling, Raymond

Awarding institution:
King's College London

KCL Department of Biophysics: PhD Thesis by Raymond Gosling 'X-ray diffraction studies of Deoxyribose Nucleic Acid'.

Abstract of Thesis submitted by R.G. Gosling for the Ph.D. Examination

A review is given of evidence on the molecular structure of desoxyribose nucleic acid (DNA). The preparation of orientated NaDNA specimens and the micro-techniques used to obtain diffraction patterns from single fibres, (50μ diameter), are described.

By varying the water content of highly orientated fibres it is shown, from the changes in X-ray diagram, that NaDNA may exist in two different molecular phases, Structure A (crystalline) and Structure B (paracrystalline); X-ray fibre diagrams published hitherto have been from mixtures of these two phases. Qualitative conclusions are drawn as to molecular configuration, inter-molecular linkages and the role of water in Structures A and B.

From measurement of reflection positions in a Structure B diagram it is shown that the molecular configuration is a helical arrangement of two polynucleotide chains, with 10 nucleotides per turn of each chain.

A procedure is described for deducing the integrated intensities from the maximum photographic intensities of the reflections in the "A" diagram. These are used to calculate the cylindrical Patterson function of Structure A, from which the configuration of the phosphorus atoms in this crystalline state are deduced. The principle lattice vectors are also revealed, so permitting the indexing of the observed reflections and hence the calculation of the three-dimensional Patterson.

The application of superposition methods to this function are discussed. A M_6 minimum function, obtained from the Patterson of those parts of Structure A related by a halving of the unit cell in c, is shown to be in agreement with the proposed molecular two-strand helical configuration of phosphorus atoms, (11 nucleotides per turn each strand), and also to indicate inter-helical linkages.

The molecular parameters deduced from this study are discussed in relation to other physical studies of NaDNA. It is concluded that the Structure B molecular configuration probably persists in solution.

(1)

X-RAY DIFFRACTION STUDIES OF DESOXYRIBOSE NUCLEIC ACID

NOTICE

GOSLING (R.G.)

Ph.D. 1954

(Biophysics)

X-RAY DIFFRACTION STUDIES OF DESOXYRIBOSE NUCLEIC ACID.

Thesis submitted for the Degree of Doctor of Philosophy in the University of London.

by

R.G. Gosling.

The Wheatstone Physics Laboratory,
King's College,
London.

February, 1954.

CONTENTS

Page

Acknowledgments

I wish to express my sincere thanks to Professor J.T. Randall, F.R.S. for the opportunity of working in the Biophysics Research Group in his Laboratory and for the constant interest he has shown in my work.

I am deeply indebted to Dr. R.E. Franklin, who has introduced me to the techniques of X-ray crystallography and with whom I have worked most closely throughout this investigation.

My thanks are due to the various members of the Department, especially Dr. M.H.F. Wilkins and Dr. A.R. Stokes, who have assisted me by discussion and technical advice.

I should also like to acknowledge the assistance that I have received from Mr. Burrows, the Technical Staff of the Physics Laboratory and the Staff of the Physics Workshop, whose ready and capable co-operation has greatly helped this work.

Finally I wish to give my best thanks to Miss F.H. Ticehurst of the Photographic Department for the preparation of many of the Plates and to my sister, Miss M.A. Gosling, for typing the script of this thesis.

INTRODUCTION

Nucleic acid was first recognised as a separate entity by Friedrich Miescher, working at Basel in 1872.[1] Since that time much work has been done in the study of single nuclear components. The importance of nucleic acid has been seen to lie in its inclusion with proteins to form the nucleoproteins, a class apart from other types of conjugated proteins. It has been recognised that it is the nuclear proteins that govern the phenomena of cellular development and the living process. More recently, following the transformation experiments on pnenmococci by McCarty and Avery [2 a.b.], evidence has accumulated which points to the nucleic acid alone in the chromatin material as the body which gives the specificity of cellular development. A knowledge of the structure of the nucleic acids would, therefore, be of interest on its own account in addition to any help it might give in studies of the nucleoproteins.

The nucleic acids which have been most completely studied are those from yeast and from calf-thymus gland. Each type is composed of phosphoric acid, one kind of sugar, and four different nitrogenous bases which may be present in varying amounts. The trebly constituted component which forms the fundamental unit is referred to as a nucleotide. Each nucleotide is the phosphoric ester of a nucleoside which is a pentose derivative of an amino-purine or a substituted pyrimidine. The nucleic acids are classified according to the pentose sugar, which is always either d-ribose or d-2'-deoxyribose. Those containing d-ribose (yeast type) are

referred to by the abreviation RNA, whilst those containing the deoxy sugar (such as the nucleic acid extracted from calf-thymus) are referred to as DNA. It is with this latter type of nucleic acid that this thesis is concerned.

DNA has been isolated from calf-thymus, spleen, liver, kidney, fish sperm and other sources of animal chromatin. Although the chemical composition of DNA, extracted from all sources, has been found to follow the general pattern outlined above, the relative amounts of the four bases varies with the species.[3] More recently it has been shown[4.5.] that an extract from a single species may be separated into fractions having different proportions of these bases.

The "molecule" of DNA is believed to consist of long unbranched polynucleotide chains, polymerised by means of the phosphate-ester bonding between the sugar rings of the nucleotides. That is, each phosphate group is linked to two nucleotides through the C_3' and C_5' positions of the sugar ring. The nucleic acid component can be separated from the protein and precipitated in the form of its sodium salt by a variety of methods. The more mild the process the greater the weight of the resulting sodium polynucleotide "particles". Molecular weights of up to 8,[6.7.8.] and 12,[9.] millions have been reported. The work to be described in this thesis is an attempt to discover something of the structure of this macromolecule. A study has been made of the X-ray diffraction patterns obtainable from orientated specimens of a high molecular weight preparation. The material used was a highly purified

preparation of the sodium salt of DNA (NaDNA) from calf-thymus, kindly given to this Laboratory by Professor Signer of Berne. The method of preparation was that published by Signer and Schwander.[6]

X-ray fibre diagrams of NaDNA have previously been obtained and discussed by Astbury et al.[10] The X-ray powder photographs of non-orientated aqueous systems of NaDNA have been studied by Riley and Oster.[11] None of these X-ray diagrams showed discrete reflections at angles corresponding to spacings smaller than 3.0A°. In the study presented here the smallest spacing observed was 2.5A°. It was realised, therefore, that even a complete quantitative study would not be capable of yielding the atomic positions in the structure. However, it was thought that this might not prove an insuperable obstacle to the determination of the general structural form, since the configuration of the constituent base components was well-established [12. 13.] and the sugar configuration suggested by analogy with RNA.[14] Furthermore, it was felt that, in any case, a comparative study of the different X-ray diagrams obtained under different conditions, combined with a detailed quantitative investigation of those diagrams which showed the highest degree of crystallinity, should yield a certain amount of new information.

One of the most striking features of NaDNA, revealed by this investigation, is the high sensitivity of the structure to the humidity of the surrounding atmosphere. It has been found that the polynucleotide chains can exist in two distinctly different helical configurations, depending chiefly on the

water content of the specimen. A general qualitative survey of this effect is described in Chapter IV. A quantitative study of the two different types of fibre diagram, corresponding to the two possible helical structures, is presented in Chapters V, VI and VII. The work described in these chapters was carried out in close collaboration with Dr. R.E. Franklin.

REFERENCES IN THE INTRODUCTION.

1. Miescher, F. (1897): Die histochemischen und physiologischen Arbeiten. Leipzig.

2(a) Avery, O.T., MacLeod, C.M. and McCarty, M. (1944) J.Exp.Med. 79 137.

(b) McCarty, M. and Avery, O.T. (1946) J.Exp.Med. 83 89,97.

3. Chargaff, E. (1950): Experientia, 6 201.

4. Chargaff, E., Crampton, C.F. and Lipshitz, R. (1953): Nature 172 289.

5. Brown, G.L. and Watson, M., (1953): Nature 172 339.

6. Signer, R. and Schwander, H. (1949): Helvet. Chim. Acta 32 854.

7. Doty, P., Bunce, B.H., and Reichmann, M.E. (1953): J. Polymer, Science 10 109.

8. Katz, S. (1952): J.Amer. Chem. Soc. 74 2238.

9. Sadron C., (1953): Informal discussion of Faraday Society. Unpublished.

10. Astbury, W.T., and Bell, F.O. (1938): Cold Spring Harb. Symp. Quant. Biol. 6 109.

Astbury, W.T., (1947): Symp. Soc. Exp. Biol. 1 66.

11. Riley, D.P., and Oster, G. (1951): Biochim et Biophys.Acta 7 526.

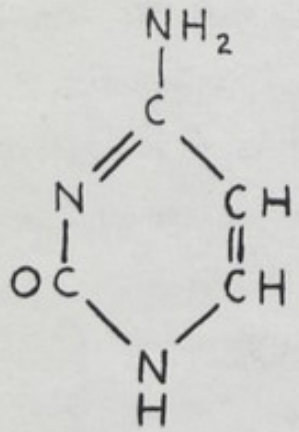
12. Broomhead, J.M. (1948): Acta. Cryst. 1 324.
(1951): Acta. Cryst. 4 92.

13. Clews, C.J.B. and Cochran, W. (1948): Acta. Cryst. 1 4.
(1949): Acta. Cryst. 2 46.

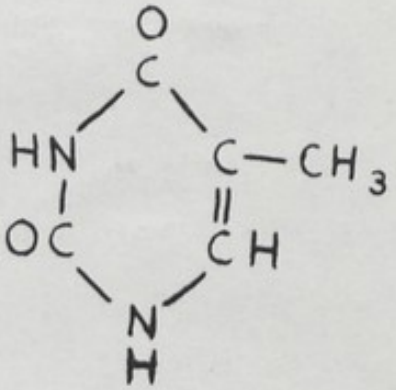
14. Furberg, S. (1949): Ph.D. Thesis. London.
(1952): Acta. Chem. Scand., 6 634.

CHAPTER I

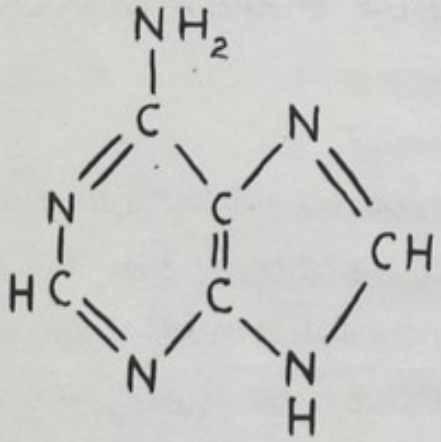
CHEMICAL EVIDENCE ON THE STRUCTURE OF DNA



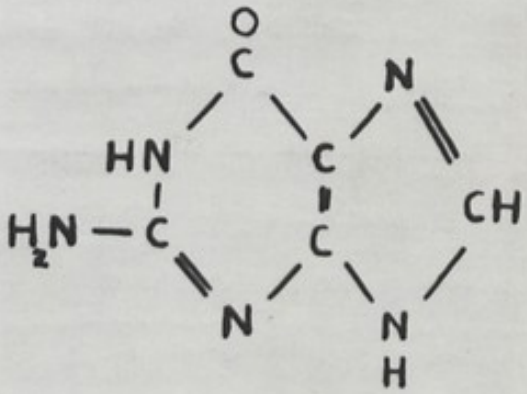
CYTOSINE



THYMINE



ADENINE



GUANINE

Fig. 1.

The polynucleotide chains of NaDNA are belived to occur mainly in an unbranched condition and to consist of phosphate di-ester groups joining β - d - 2' - deoxyribofuranose residues by 3' : 5' linkages. In calf-thymus DNA the four bases are; adenine and guanine (purines), cytosine, plus a small amount of 5 - methyl cytosine and thymine (pyrimidines). The chemical formulae of these bases are shown in Fig. 1. Since they contain groupings:-

$$-NH-C.O- \qquad -N-C.(OH)-$$

Keto Enol

and $$-NH-C.(NH)- \qquad -N-C.(NH_2)-$$

Imino Amino

both purines and pyrimidines are capable of existing in tautomeric forms. There is some evidence, [1] that the most probable tautomeric forms in the nucleic acids are the keto and amino groupings. The formulae of Fig.1. have, therefore, been drawn with these groups. Further evidence on tautomeric form is listed in Chapter II.

Structure of the Nucleosides.

The nature of the sugar in a deoxyribonucleoside was first established by Levene and Mori in 1929.[2] The sugar from the guanine nucleoside of calf-thymus nucleic acid was isolated and shewn to be a deoxypentose, and from experiments on optical rotations it was identified as d - 2' - deoxyribose. Work by Levene and his associates on RNA (1931)[1] showed that the ribonucleosides were ring nitrogen glycosides and that the sugar component occurred in the furanose condition (1932)[3].

These conclusions were confirmed convincingly in 1946 by Todd et al, [4] using synthesis techniques. It was naturally suggested that the sugar in DNA was also in the furanose form and Levene and Tipson (1935)[5] were able to show that certain derivatives of thymidine had reactions indicating a furanose structure. Makino (1935)[6] by similar tests was able to obtain the same indications from the d - 2' - deoxyriboside of guanine. However, more direct evidence was not forthcoming until the experiments of Lythgoe and Brown in 1950.[7] They showed that all four deoxyribosides, isolated from enzymic hydrolysis of herring sperm DNA, had the furanoside structure.

Experiments on the base-sugar linkage in nucleic acids were first reported by Levene and Tipson (1934)[8] who showed that the pyrimidine ribonucleosides were N_3 - glycosides. Similar methods used by Bredereck, Muller and Berger (1940)[9] showed that in thymine deoxypentoside the glycosidic radical was also attached at N_3. Spectroscopic evidence was furnished by Gulland and his associates (1938)[10] that in the purine nucleosides of both RNA and DNA the sugar was attached to the N_9 position of the purine. That this was so, for RNA, was demonstrated by direct synthesis by Todd et al. who also showed that the glycosidic centre was in the β configuration in both the purine and pyrimidine ribonucleosides, (1947, 1948)[11,12]. Thus, although since 1947 the four nucleosides of DNA have usually been formulated as 9 - β - d - 2' - deoxyribofuranosides of guanine and adenine, and the 3 - β - d - 2' - deoxyribofuranosides of cytidine and thymine, the experimental evidence with regard to the α or β nature of the glycosidic linkage is still

not conclusive and is based largely on the established structure of the ribonucleosides.

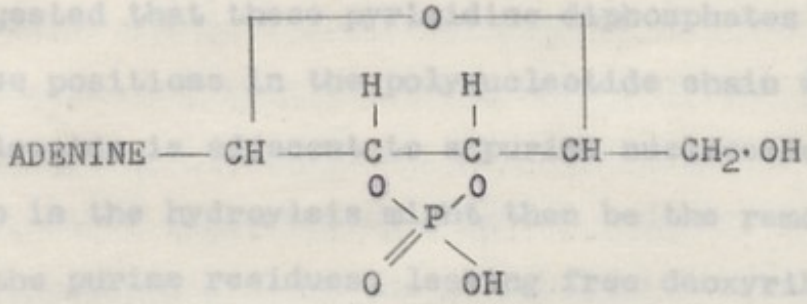
Structure of the Nucleotides.

The four nucleotides, adenylic, guanylic, cytidylic and thymidylic acid were first isolated in crystalline form by Klein and Thannhauser[13] (1933, 1934 and 1935). The existence of a fifth nucleotide, deoxy - 5 - methyl cytidylic acid, in such enzymic digests of DNA, has been demonstrated by Cohn (1951).[14] This has not yet been isolated in crystalline form. For some time it was considered that the phosphoryl group in these natural deoxyribonucleotides was in each case at C$_3'$ of the sugar residue, mainly by analogy with the natural ribonucleotides.[15, 16] In 1951 experiments with purified 5' - nucleotidase led Carter[17] to suggest that they were almost certainly 5' - deoxyribonucleotides, and this view has been confirmed in the case of thymidine by the work of Michelson and Todd. (1953).[18] They synthesised thymidine 3' - phosphate and thymidine 5' - phosphate and found that it was the synthetic 5' - phosphate which corresponded in its properties to the natural thymidylic acid, being identical in paper chromatographic and ion-exchange characteristics. The case of the other three nucleotides awaits such direct experimental evidence.

Internucleotide Linkages

The location of the phosphate residues in nucleic acids was first studied in the natural ribonucleotides by Levene and Harris (1932,[15] 1933[16]) who concluded that the phosphate residue esterified the hydroxyl at C$_3'$. However, Carter in 1950[19]

separated from a preparation of adenylic acid, two isomers a and b. These have been confirmed by Todd (1952)[20] to be the 2' and 3' phosphates of adenosine, respectively. This work has re-opened the question of the phosphate linkages in RNA. Todd[20] has shown that under acid conditions either of the isomers can produce a mixture of the two and has suggested a mechanism for phosphoryl migration involving the cyclic intermediate:-



However, Todd has pointed out that no such action can occur in DNA, because of the absence of a hydroxyl at C_2' which precludes the essential formation of the cyclic structure shown above. This, it is suggested, accounts for the greater stability of the phosphodiester linkages in DNA towards acid and alkali compared with RNA.

In DNA the C_4' in the sugar is occupied in ring formation and C_2' carries no hydroxyl group. Therefore, only the hydroxyl groups at 3' and 5' are available for internucleotide linkage. It is generally assumed that polymerisation of nucleotides occurs by means of these two linkages.

It was observed by Brown and Todd (1952)[20] that if DNA is essentially 3' : 5' linked polynucleotides then a certain amount of pyrimidine nucleoside 3' : 5' diphosphates would be expected from mild acid hydrolysis of DNA. Recently, Dekker, Michelson and Todd(1953)[21] have isolated thymidine

3' : 5' diphosphate and deoxy-cytidine 3' : 5' diphosphate from acid hydrolysis of herring sperm DNA. This was proved by showing the isolated constituents to be identical with the synthetic products prepared by phosphorylating thymidine and deoxycytidine. The authors remarked that the amount of diphosphates produced by the hydrolysis was rather high for them to have been formed solely by random fission. They suggested that these pyrimidine diphosphates could arise from those positions in the polynucleotide chain where a pyrimidine nucleoside is adjacent to a purine nucleoside. The initial step in the hydroylsis might then be the removal of some of the purine residues, leaving free deoxyribose residues attached through phosphate groups to other similar residues and to pyrimidine nucleoside residues. Then further hydrolysis might lead to rupture between the sugar residue and phosphorus, so that pyrimidine nucleosides bearing two phosphate groups would be produced.

This work therefore constitutes further proof that the macromolecule of DNA consists of 3' : 5' linked polynucleotides.

Nucleotide Sequence.

It was for some time believed that DNA contained the four bases, cytosine, thymine, adenine and guanine, in equimolar proportions, and the suggestion was made that the fundamental unit of the polynucleotide was a tetra-nucleotide. Careful analysis by Chargaff and co-workers (1949)[22]. have shown this to be untrue. Their results, together with the more recent studies of Sinsheimer and Koerner (1952)[23]. and of Markham and Smith (1952)[24] of the fragments obtained by

enzymatic digestion of DNA, have suggested that the sequence of nucleotides in the chain is complicated. This conclusion was to be expected when one remembers the biological specificity demanded of the molecule.

It was pointed out by Stern (1947)[25] that the two purine and the two pyrimidine bases possessed configurations at corresponding carbon atoms, namely NH_2 groups in adenine and cytidine and C =O groups in guanine and thymine, which were known to give rise to strong hydrogen bond formation. Since such hydrogen bridges rarely extend over distances larger than $3A^{o}$, Stern suggested that purine and pyrimidine bases formed adjacent pairs along the polynucleotide chain. Such a pairing of the bases has been postulated in the recent detailed atomic model put forward by Watson and Crick (1953)[26]. In this model the purine-pyrimidine complex is planar. It is formed by hydrogen bonding between guanine and cytosine (guanine and 5-methyl cytsoine) denoted by G : C, adenine and thymine denoted by A : T. It is further postulated that, for the spatial arrangement demanded by the model, this hydrogen bonding is specific if the bases exist in the keto form. That is, combinations other than G : C, and A : T are excluded. This idea is supported by the results of Chargaff et al, who in 1951[27] drew attention to the constancy of the 1 ; 1 ratio found for amounts of G : C and A : T in analyses of preparations from a wide variety of sources; the G : A ratio had been found to vary quite considerably. Wyatt (1952)[28] has also reported similar findings.

Recent work by Chargaff et al (1953)[29] and by Brown and Watson (1953)[30] has pointed quite startlingly to the fundamental nature of this 1 : 1, A : T, and G : C ratio. These workers have shown that a purified preparation of calf-thymus DNA, hitherto regarded as a macromolecule formed by a complex, but fixed, specific and single sequence of nucleotides, can be separated into several fractions. In these fractions the ratio of A + T to G + C, varies widely, but the A : T, and G : C, 1 : 1 ratio is maintained. Chargaff, in considering the implications of these results, concluded that "calf-thymus DNA is not one molecular unit but consists of a very large number of differently constituted individuals with an entire spectrum of chemical structural graduations."

Macromolecular Structure

The electrometric titration of NaDNA, as studied by Gulland and Jordan (1947),[31,32] revealed no groups titratable when acid or alkali was first added to the solution in water, over the range pH 5.0 to 11.0. If the pH was once taken outside these limits there occurred a rapid liberation of groups titrating in the ranges pH 2.0 - 6.0 and pH 9.0 - 12.0 respectively. On back-titrating with acid or alkali a different curve was obtained from that produced by the initial titration. Furthermore, the back-titration curve was the same whether from pH 2.5 or from pH 12.0. Gulland and Jordan concluded that the groups titrated between pH 8.0 and pH 12.0 were the hydroxyl groups of the bases whilst those titrating between pH 2.5 and pH 6.3 were the amino groups, and since acid and alkali had the

same effect in liberating these groups some linkage must have existed between them. They suggested that, in the original neutral solution, these groups were linked together by hydrogen bonds. It was further pointed out that, for steric reasons, such bonding could not occur between the amino and hydroxyl groups of the same base, but must exist between nucleotides of the same or adjacent chains.

Whilst the base groups appeared to be initially blocked, no difference in the forward and back titration of the phosphate groups was observed. Gulland and Jordon reported four primary phosphoryl, and not more, but probably lessthan, 0.25 secondary phosphoryl dissociations for every four atoms of phosphorus and concluded that the data was consistent with an unbranched chain structure. However, it was pointed out by Jordon (1952),[33] that this estimate of secondary phosphoryl dissociations was based on the ratio of amino-groups to phosphorus determined by chemical analysis, since shewn by Chargaff[22] (and by Wyatt[34]) to be too high. The correct value was 2.6 amino groups for every 4 atoms of phosphorus. This had the effect of increasing the estimated amount of secondary phosphoryl dissociation to the value obtained by Lee and Peacocke.[35]

In 1951 Lee and Peacock[35] reported results of electrometric titration of NaDNA which confirmed the suggestion of Gulland and Jordon that the difference between the forward and back-titrations was due to hydrogen bonding between the base groups in different polynucleotide chains; they also confirmed that these linkages were responsible for the high structural vicosity of DNA

solutions around neutral pH. Further evidence to support these suggestions was advanced by showing that ultrasonic irradiation decreased the displacement of the forward from the back-titration curve and at the same time reduced the structural vicosity.

Contrary to the previous work of Gulland and Jordon, but in agreement with the corrected values of that work (as explained above) Lee and Peacocke reported that the number of secondary phosphoryl groups lay in a range from just above the upper maximum suggested by Gulland and Jordon to values as high as 1 for every 10 phosphorus atoms. They pointed out that the molecular or particle weight of the preparation studied was so high that these secondary groups could not be situated solely as the ends of unbranched chains. Now the d - 2' - deoxyribofuranose residue has only the C_3' and C_5' positions available for internucleotide linkage so that branching, if it occurs at all, must take place at triply bound phosphorus atoms. Therefore in a branching structure, if there are no free sugar hydroxyl groups, the number of secondary phosphoryl dissociations should equal the number of triply bound phosphorus atoms and also the number of phosphorus atoms without corresponding base units. That is what Lee and Peacocke found; they therefore, postulated the branching structure shown schematically in Fig. 2. It is difficult to reconcile this interpretation of the data with the general body of chemical and physical evidence on DNA.

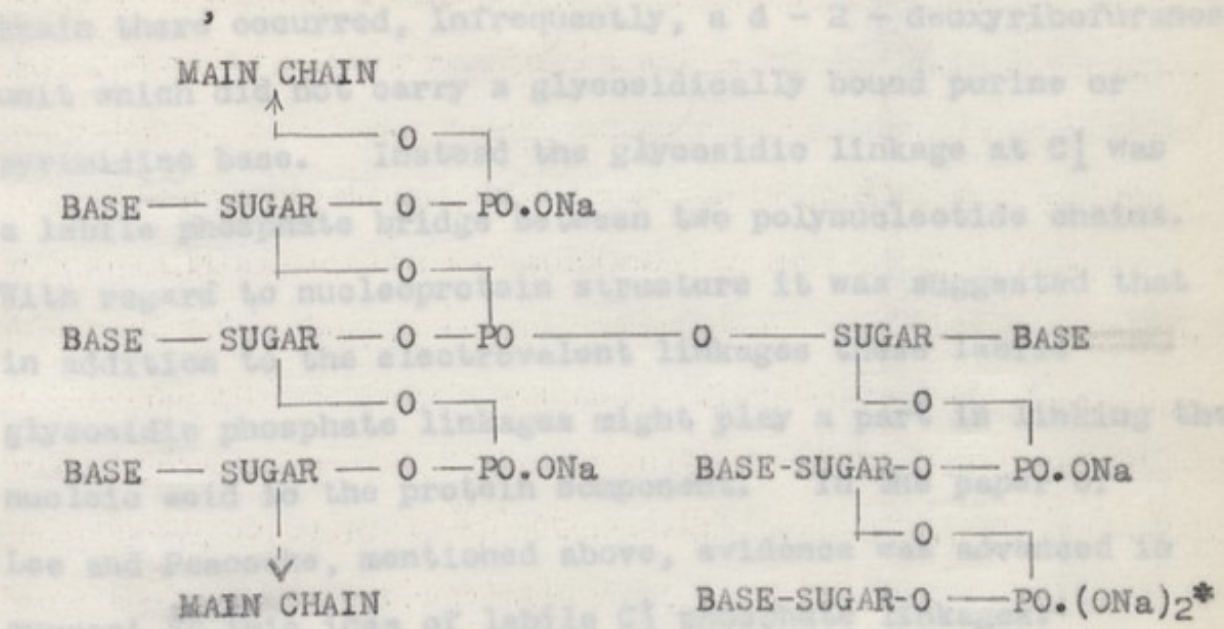


Fig. 2 A branching structure for NaDNA, suggested by Lee and Peacocke.[35]

* Secondary phosphoryl group.

Previously in 1951 Stacey et al[36] had reported work interpreted as showing that in the nucleic acid polymeric chain there occurred, infrequently, a d - 2' - deoxyribofuranose unit which did not carry a glycosidically bound purine or pyrimidine base. Instead the glycosidic linkage at C_1' was a labile phosphate bridge between two polynucleotide chains. With regard to nucleoprotein structure it was suggested that in addition to the electrovalent linkages these labile glycosidic phosphate linkages might play a part in linking the nucleic acid to the protein component. In the paper of Lee and Peacocke, mentioned above, evidence was advanced in support of this idea of labile C_1' phosphate linkages.

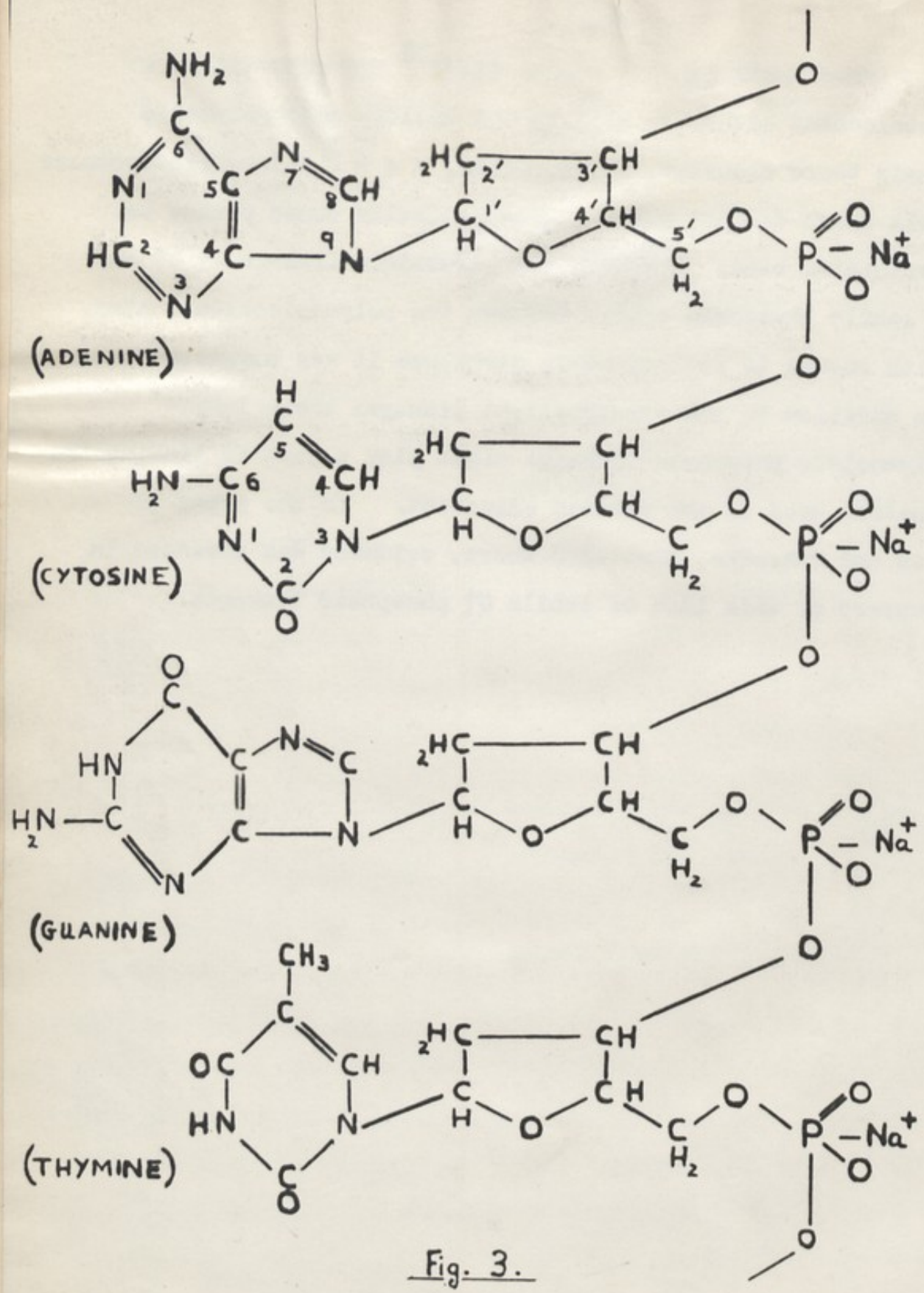


Fig. 3.

Conclusions

The evidence reviewed briefly in this Chapter establishes the chemical structure of the NaDNA polynucleotide chain. Thus a length of such a chain with the four different bases placed consecutively may be written as in Fig. 3. The titration data discussed in the last section shows that whatever the spatial configuration of this chain structure, the phosphate groups are in a readily accessible position whilst the base groups are shielded.

REFERENCES IN CHAPTER I

1. Levene, P.A. and Bass, L.W. (1931): Nucleic Acids (Chemical Catalog Company, New York).
2. Levene, P.A. and Mori, T. (1929): J.Biol. Chem, 83, 803.
3. Levene, P.A. and Tipson, R.S. (1932): J.Biol. Chem. 97, 491.
4. Todd, A.R. (1946): J.Chem. Soc. 647.
5. Levene, P.A. and Tipson, R.S. (1935): J.Biol.Chem.109, 623.
6. Makino, A. (1935): Biochem. Z. 282, 263.
7. Lythgoe, B. and Brown, D.M., (1950): J.Chem. Soc. 1990.
8. Levene, P.A. and Tipson, R.S. (1934): J.Biol.Chem. 104, 385.
9. Bredereck, H. Muller, G. and Berger, E.(1940): Ber. 73, 1058.
10. Gulland, J.M. and Storey (1938): J.Chem.Soc. 259, 262.
11. Howard, G.A. Lythgoe, B. and Todd, A.R. (1947): J.Chem.Soc. 2, 1052. (first synthesis of a natural ribonucleotide: 3 - - d - ribofuranosidecytosine).
12. Davoll, J. Lythgoe, B. and Todd, A.R. (1948): J. Chem.Soc. 967. (9 - - d - ribofuranosideoadenine).
(1948): J.Chem.Soc.1685. (9 - - d - ribofuranosidoguanine)
13. Klein, S. and Thannhauser, S.J.(1933): Z Physiol.Chem.281.173.
(1934): ibid. 224, 252
(1935): idid. 231 96.
14. Cohn, W.E.(1951): J.Amer. Chem. Soc. 73, 1539
15. Levene, P.A. and Harris, S.A. (1932): J.Biol. Chem. 98, 9.
16. Levene, P.A. and Harris, S.A. (1933): J.Biol. Chem.101, 419
17. Carter. (1951): J. Amer. Chem. Soc. 73. 1537.
18. Michelson, A.M. and Todd, A.R. (1953): J. Chem. Soc. 951.
19. Carter (1950): J.Amer Chem Soc. 72. 1466.

References in Chapter I (contd).

20. Brown, D.M. and Todd, A.R. (1952): J.Chem.Soc. 44, 52.

21. Dekker, C.A. Michelson, A.M. and Todd, A.R. (1953): J.Chem Soc. 947.

22. Chargaff, E. Visher, E. Doniger, R., Green, C. and Misani, F. (1949): J.Biol. Chem. 177. 405.

23. Sinsheimer, R.L. and Koerner, J.F. (1952): J.Amer.Chem.Soc. 74, 283.

24. Markham, R. and Smith, J.D. (1952): Nature, 170, 120.

25. Stern, K.G. (1947), Yale Journal of Biology and Medicine. 19, 944

26. Watson, J.D. and Crick, F.H.C. (1953): Nature. 171, 737.

27. Chargaff, E. (1951): Fed. Proc. 10. 654.

28. Wyatt, G.R. (1952): J.Gen. Physiol. 36. 201.

29. Chargaff, E. Crampton, C.F. and Lipshitz, R. (1953): Nature 172, 289.

30. Brown, G.L. and Watson, M. (1953): Nature 172, 339

31. Gulland, J.M. and Jordon, D.O.(1947): S.E.B. Symp.1. 56.

32. Gulland, J.M. Jordon, D.O., and Taylor, H.F.W. (1947): J. Chem. Soc. 1131.

33. Jordon, D.O. (1952): Ann.Rev. Biochem. 21, 2341.

34. Wyatt, G.R. (1951): Biochem.J. 48, 584.

35. Lee, W.A. and Peacocke, A.R. (1951): J.Chem.Soc. 3361.

36. Stacey, M. Overend, W.G., and Webb. M. (1951): J.Chem.Soc. 2450.

CHAPTER II

EVIDENCE ON THE STRUCTURE OF NaDNA FROM PHYSICAL STUDIES OF THE MACROMOLECULE AND ITS DEGRADATION PRODUCTS.

Structure of the Pyrimidine Bases

(i) Spatial Configuration

Various substituted pyrimidines have been investigated by X-ray diffraction and spectroscopic methods, but detailed structure has been determined in only a few cases. Direct studies of isolated preparations of the pyrimidines known to occur in DNA have not been reported. However, Furberg[1] in an analysis of the structure of the ribonucleoside cytidine has determined the structure of the pyrimidine cytosine.

From their studies on halogen substituted pyrimidines Clews and Cochran[2,3] concluded that the pyrimidine ring was planar.

(ii) Tautomeric Form.

Clews and Cochran suggested that the crystalline pyrimidines studied[2,3] were in the amino and not the imino form. Pitt[4], in the analysis of the structure of 2 hydroxy - 4.6 dimethyl pyrimidine, found that although the hydrogen appeared to be covalently bound to the oxygen the C - OH bond was 1.25 + 0.04A° and so possessed considerable double bond character. This indicated a tendancy towards the Keto form.

From infra-red spectral studies of crystals of cytosine, uracil and thymine, Blout and Fields[5] concluded that cytosine

was in the enol form whilst uracil and thymine were in the keto form.

The major ultra-violet absorbing configuration in both the pyrimidines and the purines was shown by Cavalieri et al.[6,7] to be the conjugated system formed by the atoms in the 3,4,5 and 6 positions of the pyrimidine ring (see fig. 3 Chapter 1). Therefore, it was to be expected that the tautomeric form would effect the ultra-violet spectrum. However, the spectrum of both the purines and pyrimidines in aqueous solution (at room temperature) has been found to be a broad band between 240 mμ. to 280 mμ. In the case of Uracil, Brown and Randall[8] have shown that this lack of specificity can be avoided by working at liquid air temperatures. However, uracil was the only one of the pyrimidies studied to show fine structure. In the recent studies by Marshall and Walker[9] of the U.V. absorption of substituted pyrimidines in solution, care was taken to adjust the pH to ensure the presence only of neutral molecules, cations or anions, so that the spectra obtained were not confused by being those of a mixture of ions and neutral molecules. Marshall and Walker concluded that the pyrimidines occurred in the keto form.

Structure of the Purine Bases.

(i) Spatial Configuration.

The first X-ray studies of a purine derivative were those by Bernal and Miss Crowfoot[10] on adenine hydrochloride. The complete structure of both the purines occurring in DNA has been determined by Miss Broomhead working on adenine and guanine hydrochlorides.[11,12]

The pyrimidine and iminazol rings were found to be co-planar and bond lengths and bond angles in both structures were very closely similar. More precise data on the adenine hydrochloride have been obtained by Cochran[13].

(ii) Tautomeric Form.

The position of the hydrogen atoms in the adenine cation has been considered by both Miss Broomhead and Cochran who conclude that it is the keto-amino form which exists in the crystalline state. The location of the hydrogen atoms in the guanine cation is not known so precisely but again the keto-amino form seems most likely.

It is of significance, with regard to the evidence presented by Gulland and Jordon for hydrogen bonding between the bases,[14] and the suggestions of Crick and Watson that such bonding is specific[15] (discussed in Chapter I), that in the crystalline state both purines and pyrimidines have been shown to form strong hydrogen bonds involving the ring nitrogens and the C = O groups. The high melting point of most of the purine and pyrimidine derivatives may be attributed to this intermoleculer hydrogen bonding.

There seems to be no conclusive infra-red or ultra-violet spectrographic data on the tautomeric form of the purines occurring in DNA.

Structure of the Nucleosides and Nucleotides.

The chemical evidence reviewed in the previous chapter, has in the case of RNA been directly confirmed by the work of Furberg who has determined the structure of cytidine by single crystal X-ray analysis[1]. No such direct studies on any

deoxynucleosides or deoxynucleotides have been reported. The two most important features of nucleoside structure, with regard to the structure of DNA are:-

(i) Relative orientation of the sugar and base residues.

(ii) The spatial configuration of the furanose ring.

In Furberg's analysis of cytidine it was shown that the $N_3 - C_1'$ bond lay in the plane of the pyrimidine ring and that it formed angles of 109° and 115° with the adjacent bonds of the d-ribofuranose ring. Therefore, contrary to the previous suggestion by Astbury[17] that the two rings were parallel, it was found that they were approximately perpendicular. From some preliminary work on uridine, adenosine and guanosine, Furberg also suggested that the mutual position of the base and sugar rings in all ribonucleosides was the same as that in cytidine. There seems to be no reason to suppose that this relative configuration of the base and sugar rings should be any different in the deoxyribonucleotides although direct confirmation of this point is necessary.

Evidence of the configuration of a furanose ring comes from the work of Furberg mentioned above and from the work of Beevers and Cochran[16]. In both these cases there was agreement that the three carbon atoms C_1' C_2' C_4' and the ring oxygen were very nearly co-planar, and that C_3' stood out of this plane by about ½ A°. But in both cases the furanose ring had hydroxyls at C_2' and C_3'. It might well be that with the hydroxyl group removed from C_2' as in DNA, the configuration of the ring is altered. This would have an immediate effect on the radius of

the helical structure of DNA proposed by Watson and Crick,[15] for this structure is quite sensitive to a change in the configuration of the sugar ring, which has been assumed to have the same spatial form as d-ribofuranose.

Blout and Fields[18] have obtained the infra-red spectra of adenosine and guanosine and assigned observed absorption frequencies to certain bonds, and have also published the infra-red spectra of adenylic, cytidylic and guanylic acids. The infra-red spectra of the four deoxynucleotides, the deoxyribosides of thymine, hypoxanathine and cytosine, the four ribonucleotides and the nucleosides together with spectra of various preparations of RNA and DNA have been reported by Fraser[19]. The greater spectral resolution obtained in this work revealed extra detail in the band spectra compared with previous data. However, it is still not possible to make definite structural determinations from such data, although it is sufficiently specific to make it very useful for analytical purposes.

Structure of the Macromolecule of DNA.

The apparent particle weight of NaDNA in solution has been measured by a variety of methods.[20. - 28.] The values obtained seem to depend on the method of extraction, and range from 10^6 to 8.10^6. The higher values are associated with the mildest of the extraction and separation methods. All determinations show the particles to be highly asymmetric in shape. The molecule of DNA is thus known to be close in form to a very long thin rod.

(i) Light Scattering.

It is possible to obtain a value of the weight average molecular weight of a substance, without any assumption as to shape of the molecule, from a knowledge of the angular distribution of the intensity of the light scattered by that substance in solution. It is necessary to extrapolate the intensity to zero angle. One of the most reliable methods of doing this is that devised by Zimm[29]. The greatest particle dimension, or related parameter, is generally obtained by comparing, graphically, the observed scattering curve with the theoretical scattering curves calculated for various ideal particle forms e.g. Rods or Coils. There is general agreement, in all the more recent work,[24 - 28] that the experimental curves for NaDNA fall between the nearest theoretical curve for a random coil and that for a stiff rod and it is suggested that this is consistent with the picture of the nucleic acid particle as a somewhat extended, partially coiled, thread. The results of Steiner[25] are shown in Table 1. These results are typical of the order of values obtained by the light scattering method for the "particle" dimensions of NaDNA samples obtained by very mild methods of extraction.

The recent work of Doty,[30] has suggested that the flexibility of the NaDNA particles in solution is dependent on the pH of the solution. Measurements on NaDNA solutions at pH 6.5, pH 3.0 and 2.6 were reported. It was found that the R value, i.e. Root mean square end to end separation of the particles, decreased on lowering the pH, whilst the molecular

TABLE I

Results from Light Scattering Measurements on two samples of calf-thymus NaDNA in 0.2M. KCl. by R.F. Steiner.[25]

Molecular Weight (± 15%)	Particle Dimension (± 10%)
$4.4.10^6$	L_1 4730A° L_2 3880A° L_3 6710A°
$7.3.10^6$	L_1 5080A° L_2 4160A° L_3 7180A°

L_1 = root mean square end to end separation calculated on the basis of monodisperse randomly kinked coils.

L_2 = root mean square end to end separation, calculated on the basis of polydisperse random coils with a ratio of weight average to number average molecular weight of 2.

L_3 = length calculated on the basis of monodisperse rods.

(The Thymonucleic acid used in these experiments was prepared by the method of Gulland, Jordan and Threlfall[35]).

TABLE II

Results of Doty et al[30] from Light Scattering Measurements on Sodium Thymonucleate in 0.2M NaCl showing the changes induced by Dilute Acid.

The following values are for the same specimen, the molecular weight remained constant throughout the pH variations, at value $7.7.10^6$.

pH.	R.
6.5	4110A°
3.0	3550A°
2.6	1760A°
6.5 (return)	4250A°

R = Root mean square end to end separation calculated for polydisperse unbranched random coils, the polydispersity corresponding to a ratio of weight to number average molecular weight of two.

The NaDNA sample was prepared by following the method of Signer and Schwander[23.(b)].

weight remained constant. (Results summarised in Table II). Furthermore this process was reversible, the R value returning to approximately the initial value on returning the pH to 6.5, there again being no change in the molecular weight. It is important to note that, as Doty points out, this work was done in 0.2 M NaCl which would mean that, at pH 3.0 and 2.6, less amino groups would have become charged than one would expect from titration data in the absence of salt. It follows that most of the inter-base hydrogen bonds, known to exist from titration work, would still be intact despite the low pH, so that a complete breakdown of the structure would not occur. However, that hydrogen bonds were broken with a resulting increase in the flexibility of the molecule, showed that these inter-base linkages play an important part in determining the configuration of the molecule. The fact that the results were reversible with pH, indicates that these hydrogen bonds can reform.

A different method of approach to the problem of interpreting the light scattering data has recently been reported by Peterlin.[31] Experimental scattering curves were compared with the scattering curves to be expected theoretically from finite lengths of a thread-like molecule with varying degrees of flexibility. The form of these structural curves was derived by following the treatment of a "thread" model of continuous mass distribution and curvature discussed by Kratky and Porod[32] in relation to X-ray scattering. The curves were plotted in terms of the parameter x, where $x = \frac{L}{a}$. "L" was the total length

of the thread and "a" was the persistence length. This latter term was defined as the average value of the projection of the straight portions of the thread on a tangent to one end of the thread. The persistence length, and hence the value x, was thus a measure of the flexibility. If a = L, x = 1 and the curve became coincident with the known scattering function for stiff rods.

In this way Peterlin has shown that his own results on a sample of calf-thymus NaDNA prepared by Signer, and the results of Bunce,[27] may be interpreted on the basis of a long coiled thread-like particle of line density 156 (Signer), 152 (Bunce). These results are shown in Table III.

In considering the results of Bunce[27] for a NaDNA solution at pH 3 Peterlin concluded that the action of acid yielded a more extended coil which was linked to neighbouring chain segments so increasing the apparent line density. This is in contradiction to the folding suggested by Doty's recent results. However, these data of Bunce, together with similar drastic effects observed by other workers are, as Doty points out,[30] most probably due to lowering the pH too rapidly. A decrease in molecular weight would then occur due to breaking of the phosphate-sugar chain at a few points along its length by local action of strong acid before uniform concentration was achieved. In the work of Doty mentioned above the pH was lowered continuously and <u>slowly</u> by dializing against NaCl solutions adjusted to the desired pH and an ionic strength of 0.2, the equilibrium pH only being reached after 6 hours.

TABLE III

Results of Peterlin[31] from calculations on Light Scattering Measurements of Sodium Thymonucleate.

Sample	Molecular Weight M	x	a	L	Line Density M/L
1	$6.7.10^6$	150	287A°	43,000A°	156
2	$4.0.10^6$	100	263A°	26,300A°	152
3	$2.64.10^6$	16	541A°	8,700A°	305

x = L/a. L = Total length of molecule.

a = Persistence length. (i.e. the average value of the projection of the straight portions of the molecule on a tangent to one end).

Sample 1 Supplied by Signer and experimental values obtained by Peterlin.

Sample 2 The results shown were calculated from the published experimental values of Bunce[27].

Sample 3 Values calculated from results of Bunce[27] on an acid degraded specimen (pH 3.0).

(ii) Sedimentation Experiments.

By measuring the sedimentation velocity of solute molecules in a solution subjected to a very high centrifugal field of known magnitude, the sedimentation constant characterising those molecules may be calculated. From this their molecular weight can be computed, if the molecular frictional resistance to sedimentation is known. It is generally assumed that this frictional coefficient is equal to the frictional resistance to diffusion, alternatively it may be obtained from visoosity data. However, when the solute molecules are highly asymmetric this assumption is only valid for very dilute solutions. This is the case for NaDNA. It is therefore, necessary to determine the variation of the sedimentation constant with concentration and extrapolate to zero concentration. This procedure inevitably involves assumption concerning the shape of the curve at very low concentrations and so introduces a certain degree of uncertainty in the final result. However, the molecular weight value so derived is without any assumptions as to molecular shape or solvation, and for this reason the method is regarded as a direct measurement.

Tennant and Vilbrant[33] have reported sedimentation studies on NaDNA in acqueous solution for concentrations up to 0.3% by weight. The NaDNA was obtained from three sources, and in each case was prepared by the method of Hammarsten.[34] The authors found that the sedimentation constant increased with decreasing concentration so rapidly as to make reasonable extrapolation to infinite dilution impossible. It was therefore assumed, in calculations of molecular weight, that

at the same finite concentration the frictional coefficients for sedimentation and diffusion were equal. By combining this sedimentation data with diffusion and partial specific volume measurements, the average cross-sectional diameters of the particles in several specimens were calculated, using the Perrin equation[35]. The values reported range from 13 to 16° A, for particle lengths of approximately 5000 to 3000A°. The highest value of the molecular weight for the specimens studied was 580.10^3, this indicated that the preparations were all somewhat degraded. However, it was remarked by the authors that all the samples studied showed in solution particles of practically the same cross-sectional diameter, regardless of their length. This, they suggested, confirmed the picture of NaDNA as a long thin column with the first step of degradation being the cutting of that column into smaller fragments of the same cross-section.

Cecil and Ogston[20] have made sedimentation and diffusion experiments in the concentration range of 0.5 to 0.03 g/100 ml. The NaDNA used was prepared in Gulland's laboratory[36]. They reported particle weights in a range from $1.0.10^6$ to $1.3.10^6$, with axial ratios for the equivalent ellipsoids between 140 and 170. These authors point out that they did not observe the very rapid increase of sedimentation constant reported by Tennant and Vilbrant[33] at low concentrations.

Steiner[25] has compared the sedimentation characteristics of nucleo-histone and NaDNA, both from calf-thymus. It was found that the limiting sedimentation constant of NaDNA was much less than that for the nucleoprotein and also that the sedimentation constant of the nucleoprotein was not so

"concentration-dependent" as that of the NaDNA. These data were interpreted as suggesting that the NaDNA particle is considerably more asymmetric than the nucleohistone particle.

(iii) Viscosity Experiments.

Intrinsic viscosity has been defined by Kraemer[37] as:-

$$\eta = \left[\frac{\ln \eta_r}{c}\right]_{c \to 0}$$

where η_r = viscosity of solution relative to that of the solvent.

C = concentration of solute.

Flory has shown[38] that it is possible to derive, from measurements of η, a value of the "viscosity average" molecular weight which under certain conditions may be reduced to a weight average value. Flory et al, have shown that the molecular configuration of high molecular weight polymers may be investigated effectively by viscosity measurements, and have derived the necessary relationships.[39, 40, 41.]

Signer, Casperson and Hammarsten[42] have made viscosity measurements on NaDNA solutions and applied formulae derived by Burgers[43]. They reported that the molecules of their preparation of calf-thymus NaDNA had, in solution, a shape factor of 300 and molecular weights ranging from 0.5 to $1.0.10^6$. However, the work of Gulland, Jordan and Creeth[44] has shown that, at all concentrations studied, the relative viscosity of their preparations of NaDNA[38] in solution decreased with the applied pressure. As these authors pointed out, this means that quantitative interpretation of viscosity data on NaDNA will be bound to in-

volve a high degree of uncertainty until the theoretical treatment of such structural viscosity data has been completed.

The qualitative interpretation of viscosity data can, however, be an important confirmation of the structural implications of other physical measurements. The change of relative viscosity of NaDNA solutions with pH. reported by Gulland et al[44] agreed closely with the titration work of those authors[14]. On the basis of the titration data it had been suggested that hydrogen bonds existed between the $N-H_2$ and C=O groups in the bases of adjacent nucleotide chains and that these bonds were broken in solution outside the pH range 5-11. In support of this, the viscosity was found to remain unchanged (at constant pressure) between pH 5.6 and 10.9, but outside these critical limits it fell to a low value. On returning the pH to 7.0 there was found to be a difference between the behaviour of the acid treated solution and the alkaline treated solution. It was reported that when a 0.243% solution of NaDNA was left at pH 12.5 for 15 minutes and then returned to pH 7.0 the relative viscosity increased steadily with time, regaining its structural character, and after 91 hours it reached a value in the region of that of the original untreated solution. For the acid treated solution, with concentrations of NaDNA up to 0.5%, there was no appreciable increase in the relative viscosity on returning to pH 7.0 after leaving the solutions at pH 3.5 for 15 minutes, but a 10% solution gelled on standing for 12 hours.

The failure of the acid treated dilute solutions to regain a high viscosity was most probably due to a too rapid lowering of the pH as discussed on page 23 . A decrease in the

molecular weight would then be likely to occur due to breaking of the phosphoric-ester linkages at certain points along the length of the polynucleotide chain by the local action of strong acid before uniform concentration was reached. Then on returning the solution to pH 7.0 although the inter-base hydrogen bonds might re-form, the particles would be much smaller, and the relative viscosity consequently less than in the original solution. The inter-nucleotide linkages being more resistant to the action of alkali there was probably very little depolymerisation on rapidly lowering the pH to 12.5; hence if the inter-base hydrogen bonding can be reformed at return to pH 7.0, particles might be formed of the same size and average configuration as were present in the original solution, the viscosity would then regain a corresponding value.

Zamenhof and Chargaff[45] have reported that calf-thymus NaDNA which was protected in the course of its isolation from the action of high temperatures, acid and alkali, as well as that of depolymerizing enzymes, gave a solution which showed no thixotropy. It is reported that this characteristic was lost irreversibly when the preparation was degraded by acid, alkali, or heat and then permitted to repolymerize; the high viscosity was then re-established but was thixotropic. It was also found that the viscosities of solutions in which the DNA had undergone repolymerization were affected by heat in quite a different manner from the original preparation. A preparation of NaDNA from yeast was reported to give results similar to those outlined above. However, it should be noted that all these experiments were made on solutions containing 0.05 M Sodium Chloride, whilst

in the experiments of Gulland and Jordon the NaDNA was in aqueous solution. Zamenhof and Chargaff have also compared the pH dependance of the viscosity of solutions of NaDNA from different sources. They found that such solutions behaved differently in their stability to acid but differed much less in their behaviour towards alkali.

That the relative viscosity of NaDNA solution was dependent on the ionic strength of the solution was first reported by Greenstein and Jenrette[46]. The addition of sodium chloride or guanidine chloride was found by Gulland et al[44] to lower the relative viscosities of solutions of NaDNA quite considerably. The viscosity decreased rapidly at first, with increasing salt concentration, and then reached a critical value at 0.01 M ~~of~~ with both sodium chloride and guanidine chloride, respectively. Increasing the concentration of salt above this critical value produced only comparatively small changes in viscosity. However, there was a rise to a small peak in 1.0M NaCL solutions, ~~f~~llowed by a decrease to the previous low value which continued to fall only very slowly. The value in 6.0M NaCL, which was the greatest salt concentration used, was little different from that in 0.01M.

This diminution of viscosity hasbeen found to run parallel with loss of streaming birefringence. Greenstein and Jenrette[46] have studied the change in birefrigence in NaDNA produced by several reagents. In some cases, e.g. with guanidine chloride, the double refraction was almost reduced to zero. These effects were found to be readily reversible, the birefringence being

restored to very nearly its original value when the added salt was removed. These data suggest that it is the folding of the molecule to a more symmetrical shape, as indicated by the loss of birefringence, that is largely responsible for the decreases in viscosity.

Zamenhof et al[47] have correlated the changes, produced by various agents, in the viscosity and bidogical activity of solutions of DNA isolated from Haemophilus influenzae; the material had a high transforming activity. This paper was one of the first reports in which any quantitative measurements of the biological activity of DNA have been made at the same time as the measurement of physical properties. The effects of heat, pH and ionic strength were found to cause the activity to vary in a similar manner to the viscosity, though at a different rate. It should be noted that the activity of a solution of ionic strength 10^{-6} dropped to 2.4% of its initial value when left to stand for only four hours at 30°C, whilst the viscosity remained at 98.5% of its previous value. Lowering the temperature to 6°C only extended the time of standing to 30 hours before a similar loss in activity was observed; the viscosity remained high (96%). A solution with ionic strength 0.02 lost 80% of its initial activity after standing 24 hours at 6°C. For the activity to be preserved an ionic strength of at least 0.1 would seem to be required. The effect of dehydration of the transforming principle on the transforming activity was also studied. The authors reported that drying under the conditions of the procedure suggested by

Signer and Schwander,[23.a.] and preservation in those conditions for 4 days at 23°C, resulted in a 99% inactivation.

Steiner[25] found that there were no structural viscosity effects in the case of nucleohistone solutions in 0.015M phosphate buffer at pH 7.2 and that the intrinsic viscosity observed lead to an axial ratio of 25. This was assuming a particle density of 1.3, g/cc and applying the equation of Simha[48] for the intrinsic viscosity of rigid prolate ellipsoids of revolution. The molecular weight of the nucleohistone was given as about 6.10^6, hence, as Steiner pointed out, the molecular asymmetry found for the nucleohistone was much less than that for nucleic acid of the same molecular weight. These data are not incompatible with the suggestion of Watson and Crick[15] that the association of protein and DNA is caused by the protein chain becoming a third strand in the DNA helix. By combining with the phosphate groups of the DNA the protein strand would remove the replusive effect of all of these charged phosphate groups. This would permit a greater folding of the particle in solution tending to a much less asymetric shape for the nucleohistone molecule.

(iv) Flow Birefringence and Ultra-Violet Dichroism.

The first measurements on the double refraction of flow in solutions of NaDNA were reported by Signer, Cassperson and Hammersten[42]. The "particles" of NaDNA were found to resemble, in shape, long thin rods and to be negatively birefringent, i.e. the greatest refractive index was at right angles to the length of the particles. The fact that the optical anisotropy was very great was interpreted as showing that the purine and

pyrimidine rings all lay in parallel planes and were nearly perpendicular to the long axis of the molecule.

The various theoretical treatments available, and the general aspects of the problem of quantitative measurements of streaming birefringence have been discussed by Edsall[49]. An approximation of the Perrin equation[35] has been applied by Edsall to the results of Signer et al,[42] and those of Kausche et al[50]; assuming a length/breadth ratio of 200, the approximate length of the particles was calculated to be 4,500A°. This corresponded to a particle breadth of 22.5A°. The associated molecular weight was between 0.5 and 1.0 million. However, the recent work of Schwander et al[51.52.] on streaming birefringence lead the authors to report a value of 8000A° and 10A°, respectively, for average particles length and breadth of calf-thymus NaDNA prepared by the method of Signer and Schwander.

Seeds has studied the ultra-violet dichroism exhibited by orientated films and fibres of NaDNA at various humidities [53a] and has discussed the implications of the measurements, including those of other workers, with regard to the molecular structure of NaDNA[53b]: Seeds concluded that the purine and pyrimidine bases were so arranged as to lie, on the average, at a small angle to the normal to the long axis of the molecule and that this angle decreased with an increase in the water content of the specimen. Observations on calf-thymus nucleoprotein were interpreted as showing that the ultra-violet absorbing groups in the structure were orientated in a manner similar to the arrangement in NaDNA.

(v) X-Ray Diffraction.

Orientated Specimens:

The first studies of X-ray diffraction patterns from orientated NaDNA specimens were those reported by Astbury et al.[54.17] Interpretation of the X-ray fibre diagrams led Astbury to suggest that the molecule of NaDNA was a "stiff column" of nucleotides piled one on top of the other in such a way that the arrangement was repeated every eight nucleotides, along the length of the column. Further, the nucleotides were considered to be "flat plates" lying at right angles to the length of the column, the effective thickness of each nucleotide being about 3.4A°. This latter suggestion required the plane of the sugar ring to lie parallel to that of the base ring system. Thus a column of such nucleotides would have a strong repeat of about 3.4A° along the axis of the column. In this way the very strong meridional arc of 3.34A°, seen in the X-ray fibre diagram, could be explained.

Astbury's interpretation of the X-ray data, as outlined above, differs from that which will be advanced here. This is because Astbury's X-ray diagrams, although interpretated as from a single phase, were due to a mixture of two different structural phases of NaDNA. This will be discussed in Chapter IV.

The work of Furberg[1] made Astbury's suggestion of a planar nucleotide rather doubtful. Furberg found that the plane of the sugar ring was almost at right angles to that of the base component in each of the four ribonucleosides. So far no structural work on any desoxyribonucleosides has been reported. However, Furberg considered that it was likely that the relative

configuration of the sugar and base rings in the DNA nucleotides would be the same as that found for the RNA components. Making this assumption Furberg suggested two possible structures for DNA. These are shown in Fig.1. In each structure the base rings were co-planer and 3.4A° apart in the direction of the long axis of the molecule. The phosphorus atoms also lay close to this plane. Structure I was a helical array of phosphorus atoms surrounding the sugar rings and base groups, which were turned inwards toward the helical axis. The diameter of this single strand helix was about 13A°. The pitch of the helix was given as about 27A° to correspond with the layer line spacing reported by Astbury. Structure II was a rather open arrangement consisting of a single zig-zag chain of phosphate groups, with the base rings projecting on alternate sides of the phosphate chain. The total breadth of this structure was given as 18A°.

Furberg pointed out that although many possible structural forms between those of 1 and II might exist, Structure I was to be preferred due to the stability given by the van der Waals attraction between the base rings, which formed a flat pile along the helical axis. Also the density of Structure I was higher than that of Structure II.

The first workers to suggest a helical structure for the NaDNA molecule, from a direct study of the X-ray fibre diagram, were Stokes and Wilkins in 1951.[55] Since the completion of the experimental work presented in this thesis, the micro-techniques used by Miss Franklin and the author[56] have been extended by Wilkins et al,[57] who have obtained X-ray fibre diagrams of NaDNA showing nearly twice the number of separate reflections observed

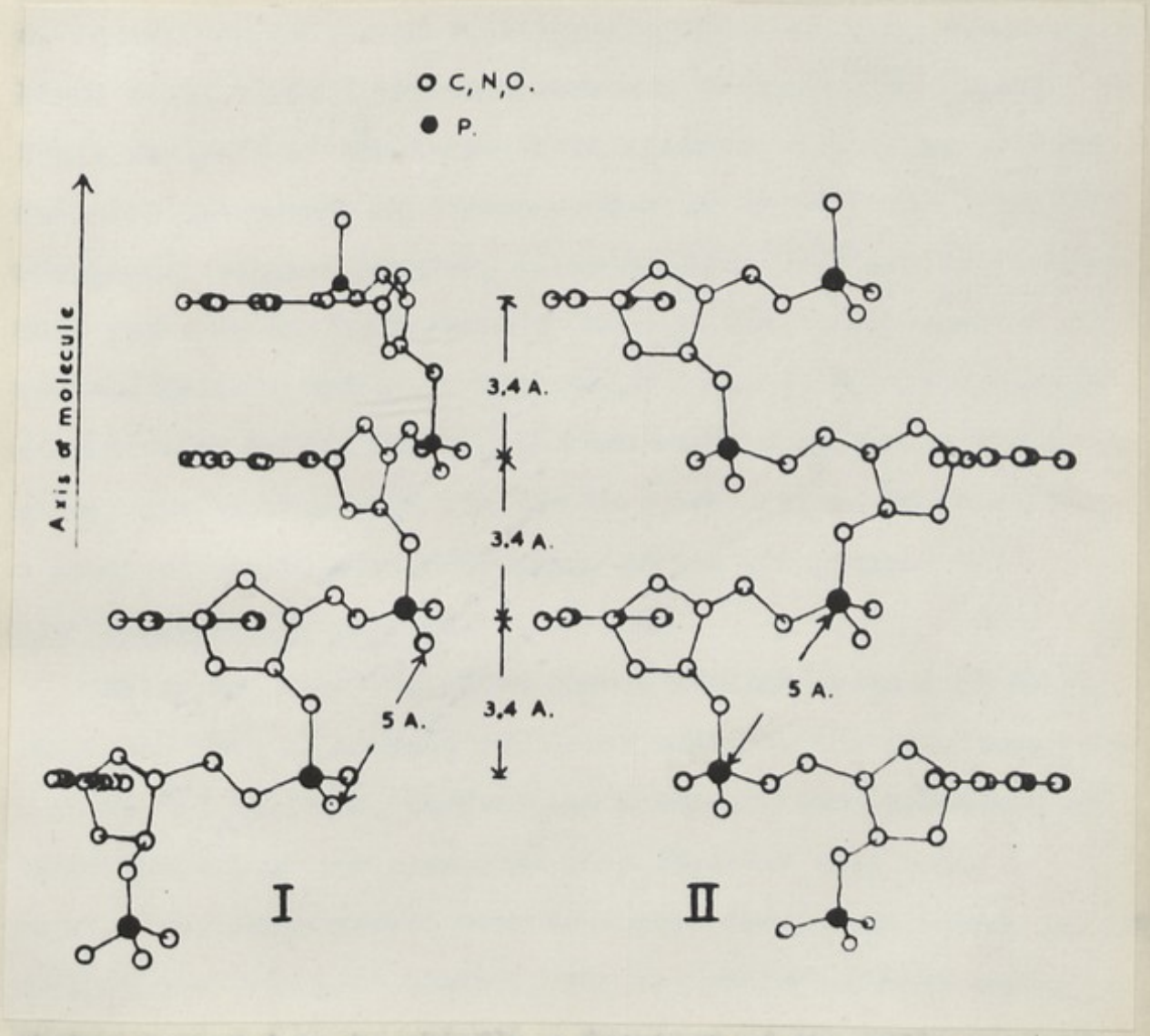


FIG. 1 Two possible molecular structures for DNA; suggested by Furberg

on the "crystalline" patterns discussed here. These authors have also shown[57] that the same "crystalline" diffraction pattern can be obtained from NaDNA prepared from several different sources; e.g. Calf-thymus, Mouse sarcoma, Human white blood cells, E.Coli, Pneumococcus and Paracentrotus sperm. In their analysis of the X-ray fibre diagrams Wilkins et al have followed the method of direct comparison between the weighted reciprocal lattice diagram, obtained from the observed diffraction pattern, and that shown by calculation to be expected from various helical models. The conclusions, as to the molecular structure of NaDNA, which have been reached so far in their study, are in many ways similar to those arrived at here from a study of the Patterson diagrams of the structure.

Non-Orientated Specimens.

Riley and Oster[58] studied non-orientated systems of NaDNA and found that at certain low water contents the specimens gave a sharp, crystalline, diffraction pattern. On increasing the water content of the specimens they observed that this crystalline state passed over to a more disordered state, termed by them the "wet gel" state. The transition between these states appeared to be chiefly a function of the water content of the specimen and was reversible. They interpreted their data in terms of inter-molecular changes due to micelle formation, or alternatively, as due to a continuous intra-molecular change. The relation of that work to the work presented here is discussed in Chapter IV.

(vi) Electron Microscopy.

One of the first reports of electron microscope studies, of

NaDNA specimens, was that by Scott[59] in 1948. The material used was isolated from rat liver following the methods of Mirsky and Pollister.[60] Micrographs, taken with an R.C.A. model EMB electron microscope, were obtained from drops of solutions with various concentrations, all of the order of 10^{-6} g/cc. These specimens were shown to contain a net work of laterally branching fibres of variable length and width. No values were given by Scott for the diameters of these fibres, but from the published micrographs they seemed to vary from about 400A° down to the limit of resolution, which appeared to be 50A°; the lengths of the fibres commonly exceeded 3 μ .

Bayley[61] has obtained micrographs of surface replicas of NaDNA, which appeared to exhibit globules of about 100A° diameter. It was suggested by Bayley that these globules were coiled up fibrils. Milward[62] has taken micrographs of gels of NaDNA spread directly across the microscope grids, without the support of collodion film.

In the studies mentioned above the "particle" dimensions observed were undoubtly those of multiple filaments. The first electron micrographs to show particles of comparable size with the molecular dimensions obtained from other methods, was that of Williams.[63] The specimens used were obtained by means of a particular freeze-drying technique, developed for use with biological material in an attempt to preserve any three dimensional structure that might be present. Micro-volumes of a water solution of 0.01% NaDNA, (prepared by the method of Signer and Schwander[23]), were sprayed onto microscope grids at

liquid air temperatures, the ice was then sublimed off under vacuo. Micrographs of portions of such freeze dried droplets, (average droplet volume was estimated at 10^{-9}ml), showed a net-like three dimensional structure of fibrils. In many places these fibrils appeared to have aggregated into bundles, it was suggested that this might have happened during vacuum sublimation. Since the individual fibrils were of very small diameter the width of accumulated "shadowing" metal was greater than that of a fibril, therefore measurements were made of the shadow lengths and the height of the fibrils deduced from the known shadow angle. In this way the fibrils were estimated to have diameters of approximately 15A°. Similar results to those of Williams have been reported by Kahler et al.[64. a.b.] These authors obtained micrographs of NaDNA showing long fibrils, of 10 - 20A° diameter, which appeared to have a strong tendency to intertwine with neighbouring fibrils.

Conclusions

The experimental data that have been accumulated by measurements of the physical properties of DNA are often conflicting, it is therefore, difficult to draw from such data more than very general conclusions as to the structure of the macromolecule. That the molecule was close in shape to a very long, thin, flexible thread has been suggested by light scattering, streaming birefringence and ultracentrifuge data. This suggestion would now appear to be certain in view of the recent electron microscope work. Measurements of the Ultra-violet dichroism, exhibited by orientated specimens of NaDNA, have shown that the base rings lie in planes parallel to each other and inclined at an angle between 60 - 90° to the long axis of the molecule, also that the polynucleotide chains are so arranged that if the water content of the specimen is increased this angle of inclination of the bases approaches more closely to 90°. Information as to the detailed arrangement of the polynucleotide chain, or chains, in this long thread-like molecule has so far only been obtainable from a study of the X-ray diffraction patterns given by NaDNA. The suggestions of previous workers have been outlined above and the structural implications of the study presented here are discussed in the following Chapters.

REFERENCES IN CHAPTER II.

1. Furberg, S. (1949) Ph.D. Thesis. London.
2. Clews, C.J.B., and Cochran, W. (1948), Acta.Cryst. 1. 4.
3. Clews, C.J.B., and Cochran, W. (1949), Acta.Cryst. 2. 46.
4. Pitt, G.J., (1948) Acta. Cryst. 1. 168
5. Blout, E.R., and Fields, M. (1950).J.Amer.Chem.Soc.72. 479
6. Cavalieri,L.F., Bendich, A., Tinker, J.F. and Brown, G.B. (1948), J.Amer.Chem.Soc. 70. 3875.
7. Cavalieri, L.F., and Bendich, A. (1950). J.Amer.Chem.Soc. 72. 2587.
8. Brown, G.L., and Randall, J.T., (1949), Nature.163. 209.
9. Marshall, J.R., and Walker, J. (1951). J.Chem.Soc. 1004.
10. Bernal,J.O., and Crowfoot, D. (1934). Nature. 131. 911.
11. Broomhead, J.M., (1948). Acta. Cryst. 1. 324.
12. Broomhead, J.M., (1951). Acta. Cryst. 4. 92.
13. Cochran, W. (1951). Acta. Cryst. 4. 81.
14. Gulland, J.M. and Jordan, D.O. (1947).S.E.B. Symp. 1.56.
15. Watson, J.D. and Crick, F.H.C. (1953). Nature 171. 737.
16. Beevers, C.A. and Cochran, W., (1947), Proc.Roy.Soc.A. 190. 257.
17. Astbury, W.T. (1947). S.E.B. Symp. 1. 66.
18. Blout, E.R. and Fields, M. (1949).J.Biol.Chem.178. 431.
19. Fraser. D.R.B.(1951). Ph.D. Thesis. London.
20. Cecil, R. and Ogston, A.G., (1948). J.Chem.Soc.1382.
21. Ogston,A.G., (1950). Trans.Faraday Soc.46. 791.
22. Kahler, H.(1948). J. Phys. Colloid. Chem. 52. 675.

References Cont.

23. Signer R. and Schwander, H. (a) (1949). Helvet.Chim Acta . 32. 854. (b) (1950). Helvet.Chim.Acta. 33. 1522.

24. Katz, S. (1948). J. Amer. Chem. Soc. 74. 2238.

25. Steiner, R.F. (1952). U.S.A. N. M.R.I. Memorandum Report 52 - 2 10. 947.

26. Doty, P., and Bunce, B.H. (1952). J. Amer.Chem.Soc.74. 5029.

27. Bunce, B.H. (1951) Thesis. Cambridge. Mass. U.S.A.

28. Rowen. J.W., (1953). Biochimica et Biophysica Acta.10. 391.

29. Zimm, B.H. (1948) J.Chem. Phys. 15. 1093.

30. Reichmann, M.E. Bunce, B.H., and Doty, P. (1953). J. Polymer. Science. 10. 109.

31. Peterlin, A. (1953). Nature. 171. 259.
(1953). J. Polymer. Sc. 10. 425.

32. Kratky, O., and Porod, G. (1949). Proc.Int.Coil. Macromol. Amsterdam. 250.

33. Tennent, H.G., and Vilbrandt,C.F., (1943).J.Amer.Chem.Soc. 65. 424.

34. Hammarsten, E.(1924). Biochem. Z. 124. 383.

35. Perrin, F. (1936). J.Phy. Radium. 7. 1.

36. Gulland, J.M., Jordon, D.O., and Threlfall, C.J. (1947), J.Chem. Soc. 1129.

37. Kraemer, E.O. (1938). Ind.Eng. Chem. 30. 1200.

38. Flory, P.J. (1943). J. Amer. Chem. Soc. 65. 372.

39. Flory, P.J. (1949). J.Chem.Phys. 17. 303.

40. Fox, T.G. and Flory,P.J.(1949).J.Phys.and Colloid.Chem.53.197

41. Krigbaum,W.R. Madelkern,L. and Flory,P.J.(1952). J.Polymer. Sc. 9. 381.

References Cont.

42. Signer, R., Casperson, and Hammarsten, E., (1938). Nature. 141. 122.

43. Burgers, J.M.,(1937). Second Report on Viscosity and Plasticity. Verhand Nederl Kon. Akademi. v. Wetenshappen te Amsterdam Afd. Natuurkunde.

44. Gulland, J.M., Jordon, D.O., and Creeth, D.M., (1947) J. Chem. Soc. 1141.

45. Zamenhof, S., and Chargaff, E., (1950) J. Biol. Chem. 186. 207.

46. Greenstein, J.P., and Jenrette, W.V., (1940). J.Nat.Cancer. Inst. 1. 77.

47. Zamenhof, S., Alexander, H.E. and Leidy, G. (1953). J.Expt. Med. 98. 373.

48. Simha, R., (1940) J. Phys. Chem. 44. 25.

49. Edsall, J.T., (1942) Adv. in Colloid, Sc. 1. 269.

50. Kausche, G.A., Guggisberg, H., and Wissler, A., (1939) Naturwissenschaften. 27. 303.

51. Schwander, H., and Cerf. R., (1949) Helv. Chim. Acta. 32. 2356.

52. Schwander, H., (1950) J. Chim. Phys. 47. 718.

53.a Seeds, W.E. (1951) Ph.D. Thesis. London.

b Seeds, W.E., (1953) Prog. in Biophys and Biophysical Chemistry. 3. 27.

54. Astbury, W.T. and Bell, F.O., (1938) Cold Spring Harbour, Symp. Quant. Biol. 6. 109.

55. Stokes, A.R. and Wilkins, M.H.F. (1951). Cavendish Lab. Protein Conference. Unpublished.

References Cont.

56. Franklin, R.E., and Gosling, R.G., (1953) Acta. Cryst. 6. 673. 678.

57. Wilkins, M.H.F., Seeds, W.E., Stokes, A.R., and Wilson, H.R. (1953) Nature. 172. 759.

58. Riley, D.P., and Oster, G., (1951) Biochim. et Biophys. Acta. 7. 526.

59. Scott, J.F., (1948) Biochim. Biophys. Acta. 2. 1.

60. Mirsky, A.E., and Pollister, A.W., (1946) J.Gen. Physiol. 30. 101.

61. Bayley, S.T., (1951) Nature 168. 470.

62. Milward, J.L., (1953) Biochim. Biophys. Acta. 10. 5.

63. Williams, R.C., (1952). Biochim. Biophys. Acta. 9. 237.

64a Rowen, J.W., Eden, M., and Kahler, H., (1953) Biochim. Biophys. Acta. 10. 89.

b Kahler, H., and Lloyd, B.J., (1953) Biochim. Biophys. Acta. 10. 355.

CHAPTER III.

PRELIMINARY INVESTIGATIONS AND APPARATUS.

Preparation of Orientated Specimens of NaDNA.

The material used throughout this work was highly purified sodium nucleate from calf-thymus gland, kindly supplied to this Laboratory by Professor Signer and prepared according to the published method of Signer and Schwander.[1.2.] The NaDNA from this preparation is a white fibrous solid which is capable of taking up large quantities of water to form a gel, the process being a continuous one ending ultimately in solution.

(i) Sheets.

By means of suitable mechanical treatment and drying, the gel can be used to obtain highly orientated specimens. If the gel is sheared between glass plates, thin "sheets" of the material can be obtained. The molecules are orientated in the plane of the sheet and in the direction of shear. These "sheet" specimens show the highest degree of orientation when diffracting in the "edge on" position. A narrow strip of the sheet is placed so that the X-ray beam is perpendicular to the direction of shear, but lies in the plane of the sheet. The effective dis-orientation of the particles, (individual molecules or crystallites), is then the projection of their angular displacement on the direction of shear.

(ii) Fibres.

In considering the problem of obtaining more highly orientated specimens, Wilkins (1949)[3], noticed that it was possible to draw fibres from a gel of the Signer material, but not from other preparations available at that time. This was done by adding sufficient distilled water to a small piece of the fibrous solid to form a stiff gel, and then placing a

needle point, or any suitably shaped object, in the gel and slowly withdrawing it either by hand or by using the rack and pinion motion of a microscope. The latter method is to be preferred if fibres of fairly uniform cross-section are required. By suitably varying the speed of withdrawal, and also the water content of the gel, fibres of diameter from about 100μ down to less than 1μ can be obtained at will. These fibres exhibit considerable negative birefringence (Wilkins et al 1951[4]) which may be taken, to some extent, as a measure of the degree of particle orientation. It was confirmed later, when the diffraction pattern of single fibres could be obtained, that the most highly orientated specimens were those which showed complete lengths of the fibre with the same degree of birefringence, rather than different domains of birefringence at the same cross-sectional level along the fibre; the greater the thickness of the fibre the more usual is this latter appearance. Although some very highly orientated fibre diagrams have been obtained with thick fibres (50μ), in general the smaller the diameter of the fibre the greater the degree of orientation.

Initial Apparatus.

The first diffraction patterns of NaDNA, obtained by the author, were taken with a 3 cm radius Unicam single-crystal camera and copper radiation from a Raymax tube. The specimens were thin, negatively birefringent, sheets prepared by Wilkins. With such weakly diffracting material it was necessary to minimise the effect of air scattering. This was done by passing a continuous stream of hydrogen through the camera.

The only serious escape of gas from the Unicam camera occurred at the collimator entry hole. This was almost eliminated by tying a brass sleeve into a central hole cut in a circle of some very thin rubber, the outer edges of which were then waxed down on the outside of the camera. The collimator was then inserted into the camera through the tightly fitting brass sleave.

Diffraction Patterns of Sheet Specimens.

Sheet specimens were examined for signs of double orientation, i.e. preferential orientation of the particles in the plane of the sheet. However, no difference could be discerned between the diffraction patterns obtained with the plane of the specimen normal and "edge on" to the X-ray beam. In each case the beam was at right angles to the direction of shear. The pattern from a sheet specimen of Signer NaDNA in hydrogen at 90% relative humidity is shown in Plate.1. The specimen was maintained at this high humidity by passing large bubbles of hydrogen through a small quantity of water before allowing the gas to enter the camera. The relative humidity of the hydrogen was measured, before it entered the camera, by an Ednay Paper Hygrometer. The definition of the pattern was found to be best at high humidities but no systematic study of the variations of these "sheet" patterns with humidity was made.

The pattern, (Plate 2), obtained from a sheet specimen of NaDNA from herring sperm (prepared by Dr. Mary Fraser in this laboratory) at 90% R.H., was found to be similar to that of calf-thymus NaDNA at the same humidity. Both patterns show a rather diffuse reflection on the meridian, the centre of which corresponds to 6.9A°, in addition to the usual

PLATE 1 Sheet specimen of calf-thymus NaDNA. X-rays perpendicular to plane of Sheet. Unicam camera filled with hydrogen at 90% Relative Humidity. Copper Radiation. Exposure 4 hours

PLATE 2 Sheet specimen of herring sperm NaDNA. X-rays in plane of sheet. 90% Relative Humidity. Exposure 5 hours

N.B. In each case direction of shear in sheet is meridional.

3.4A° arc. These "sheet" patterns differ from those of Astbury[5], probably due to the high humidity and the method of preparation (by shearing the gel) of the specimens. Astbury et al used thicker specimens, in air at room humidities. The patterns shown here are closely related to the patterns found for single fibres at high humidity which will be discussed in Chapter IV.

"Crystalline" Diffraction Pattern from Fibre Specimens.

Owing to the large specimen-film distance and beam size of the Unicam - Raymax arrangement, it was necessary to use a bundle of NaDNA fibres as a specimen. The exposure time for a single fibre with that arrangement would have been prohibitive. Accordingly a bundle of about 30 fibres of diameters 10-30 μ was made up on a wire support. The distance between the legs of the support was made quite small, about 0.5 c.m., and long lengths of fibre wound round the support and held in position by Lepages cement, a fast drying cellulose glue. In this way the required thickness of the composite specimen was readily achieved. Suitably long fibres were obtained using a microscope rack and pinion movement to control the drawing of the fibre from the gel. In this way fibres up to 7 or 8 cm in length were drawn. Once in position on the frame, the fibres were grouped into parallel alignment by cementing them together for some distance along their lengths, from each end. The legs of the frame were also opened slightly to add tension to the bundle. However, it was difficult to obtain the fibres in perfectly parallel alignment and this was a defect of the multi-fibre specimen technique.

In order to increase the resolution, of the X-ray patterns, Cobalt radiation was used; Plate 3 shows the pattern obtained from a bundle of 30 to 40 fibres at about 90% R.H. The pattern closely resembles that to be expected from the rotation of a single crystal. Therefore a fibre must contain many crystallites aligned closely parallel to the fibre axis, but having random azimuthal orientation around the axis. The diffraction broadening in arcs about the origin of the pattern is due to the fact that the crystallites are not perfectly aligned with respect to the fibre axis.

Crystallite Size.

The breadth of single reflections measured along a line joining the reflection to the origin, is approximately constant over the whole photograph. The diffraction breadth due to the individual crystallites is therefore small compared with the geometrical broadening. Hence the effective crystal diameter cannot be less than approximately 1000A° (10^{-5}cm).

Reversible role of Water.

The high sensitivity of the NaDNA structure to the humidity of the surrounding atmosphere is shown by comparing Plate 3 (90% R.H.) with the pattern obtained from the same specimen at 15% R.H., Plate 4. It will be noted that in the latter case the three dimensional order in the specimen has decreased.

To determine the effect of strong drying, the specimen was placed in a P_2O_5 desiccator for several days and then heated in the camera, in an atmosphere of dry hydrogen to a temperature of about 80°C., by inverting a small, steam heated, copper can over the specimen. Two co-axial brass tubes were soldered to

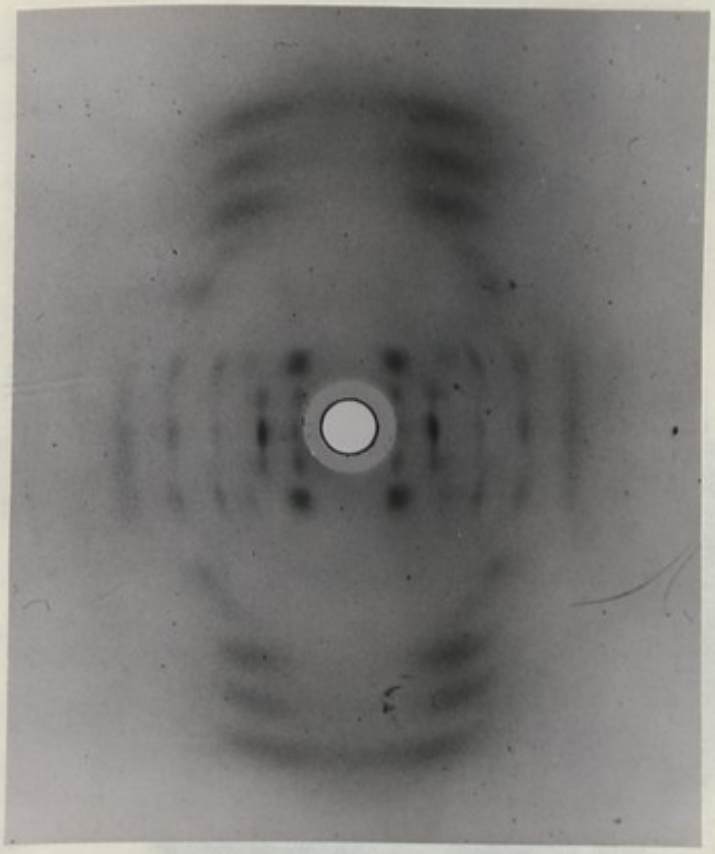


PLATE 3 Multi-fibre specimen containing about 35 fibres of NaDNA; diameters between 10 and 30μ.
Unicam camera hydrogen filled; 90% R.H.
Exposure 4 hours. Cobalt Radiation.

PLATE 4 Specimen as for Plate 3; Relative Humidity 15%
Unicam camera, hydrogen filled.
Copper Radiation.
Exposure 4 hours

the base of this can, so that steam could be passed down the inner tube and led away through the larger outer tube. These tubes also served as supports to suspend the can from the lid of the camera. A control experiment with a mercury-in-glass thermometer, placed with its bulb in the centre of the can, indicated a temperature of 78°C; the can was not shielded by the rest of the camera. The specimen was heated for two hours before the exposure, the copper can was then withdrawn to the top of the camera and maintained at steam heat throughout the exposure. The photograph obtained, Plate 5, shows that if all the water is removed and the fibres are heated, the three dimensional order disappears completely. Furthermore, large irregular holes are formed in the structure as indicated by the strong low angle scattering.

After this experiment the specimen was maintained in a hydrogen atmosphere of 90% R.H. for twelve hours. The diffraction pattern then given by the specimen, Plate 6, was the most highly orientated pattern obtained with this Unicam-Raymax-multifibre specimen technique. These experiments indicate that water is essential to the structure of the crystallites; also that the process of crystallisation is not prevented, or apparently altered, by previous strong drying or heat treatment to about 80°C.

Crystallite Formation.

From the work of Riley and Oster (Chapter II page 36) it is likely that the crystallites are formed mainly in the gel and aligned by the fibre, rather than being formed as a result of the strong orientation created by drawing the fibre. In

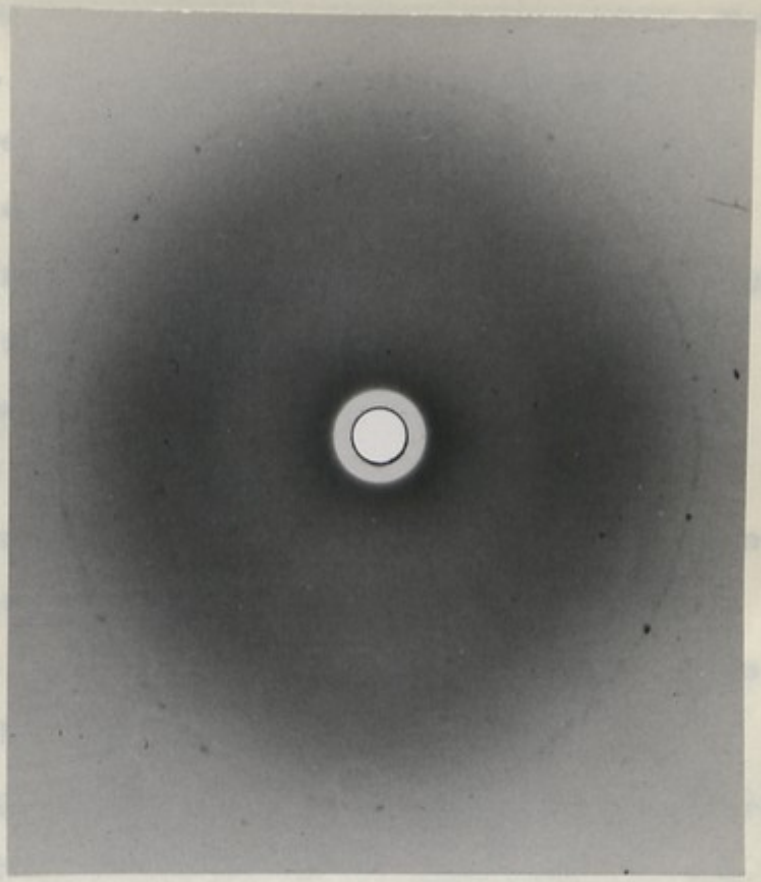


PLATE 5

Specimen as for Plate 3.

Dried over P_2O_5 and heated in camera to about 80°C prior to exposure.

Camera filled with dry hydrogen.

Cobalt Radiation

Exposure 5 hours

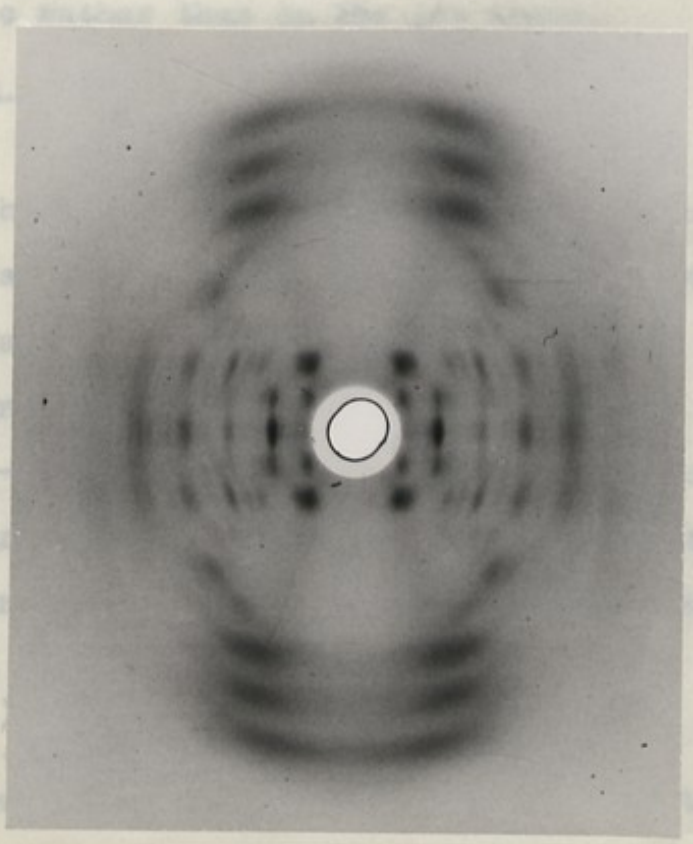


PLATE 6

Specimen as for Plate 5 (3 and 4)

Exposure taken after drying experiment shown above.

Specimen maintained at 90% R.H. in hydrogen for 12 hours prior to exposure.

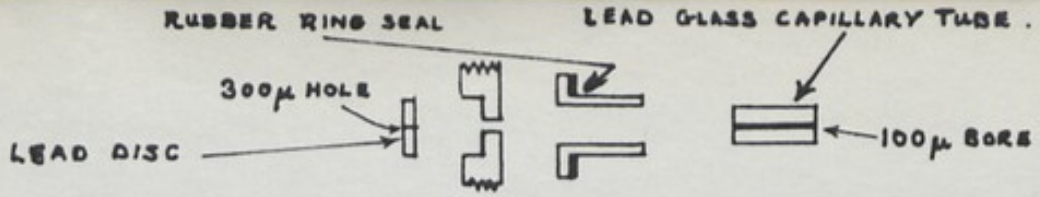
Cobalt Radiation.

Exposure 3 hours

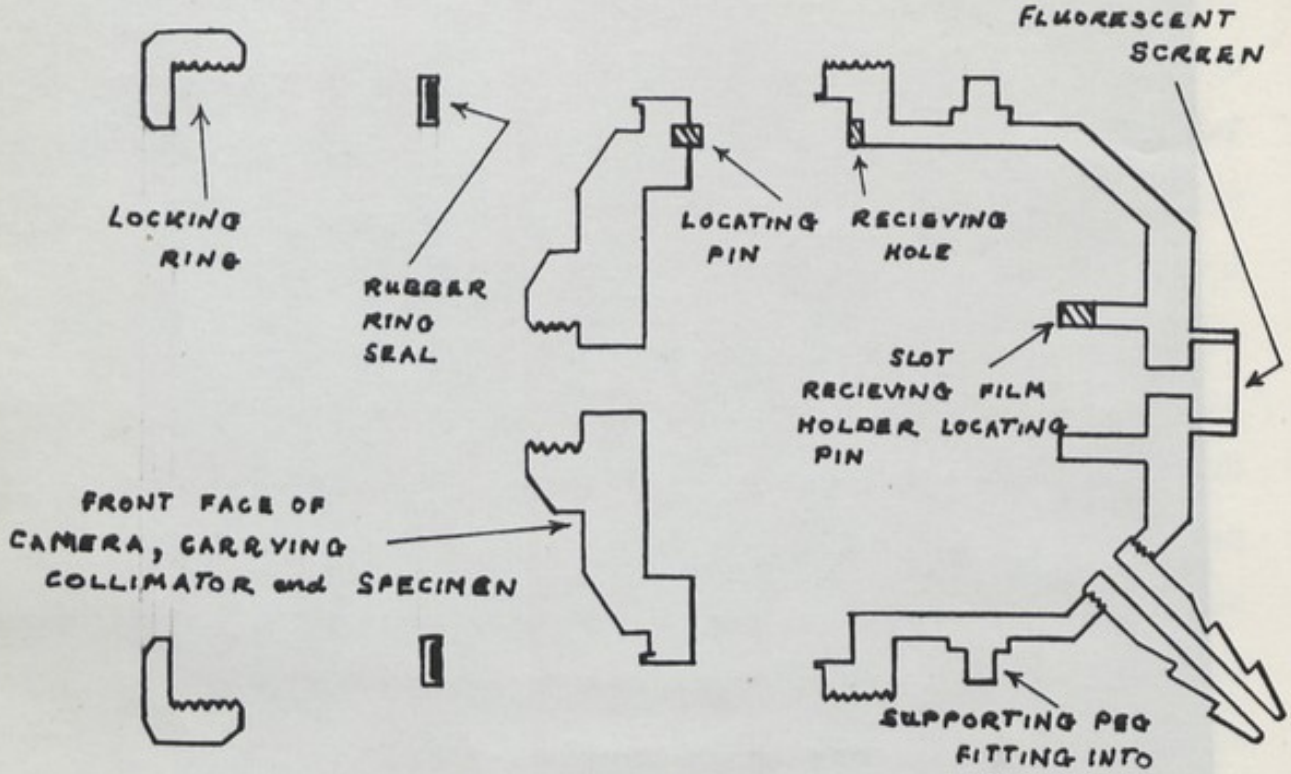
either case, since the macro-molecule is known to be a long thin chain, it is likely that one chain passes through at least two crystallites. Therefore a fibre most probably consists of a large number of regions of high three dimensional regularity connected by an irregular arrangement of inter-crystallite chain lengths, similar to the structure of rubber suggested by Bunn.[6] Some of the water up-take of these fibres will then be due to this amorphous matrix of inter-crystallite chains. The crystallites themselves would seem to require water as an essential part of their structure, which is broken down by strong drying or heating. The greater detail seen in the diffraction pattern obtained from fibres which had been allowed to take up water after strong drying, may be due to the formation of larger crystallites since re-crystallisation had to take place in the fibre rather than in the gel state.

Unit Cell.

The projection of the unit cell in a plane at right angles to the fibre axis was found to be very close to hexagonal, but it was impossible to index the rest of the photograph on that basis. However, a close agreement was found between the observed spacings, and those calculated on the basis of a C face-centered monoclinic cell, with C parallel to the fibre axis. Unfortunately, due to the small number of reflections, the layer lines for $L > 0$ proved difficult to index unambiguously with regard to the value of C^* and β^*. The value of $C^* \sin\beta^*$ i.e. λ/c, was obtained directly from the layer line period, the values of a^* and b^* were obtained from the equator. A rectangular $a^* - b^*$ net was then constructed and the various layer

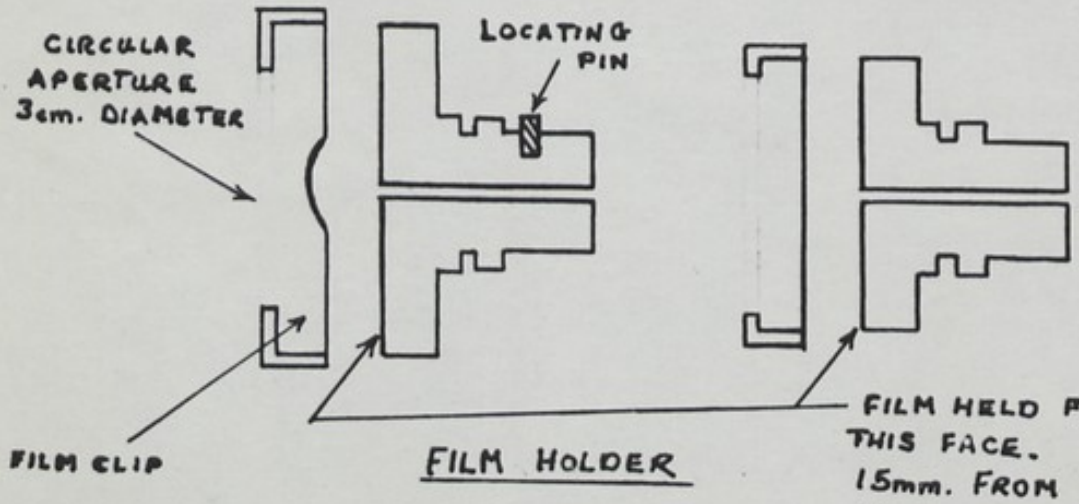


COLLIMATOR ASSEMBLEY



BODY OF CAMERA

FIG 1



FILM HOLDER

FULL SIZE SCALE DRAWING OF NORTH AMERICAN PHILIPS MICRO-CAMERA

line ξ values were indexed by graphical agreement with the net, using displacements of the origin by $l.c^* \cos\beta^*$ along the negative direction of the a^* axis for hkl values and in the positive direction for $hk\bar{l}$ values.

Indexing Plate 6 by direct inspection, in the manner outlined above, led to the values; $\lambda/a^* = 22.6A^o$, $b = 42.0A^o$, $c = 28.5A^o$. Due to the ambiguity with regard to the best choice of $C^* \cos\beta^*$, the value of β was uncertain.

The need for a Micro-Technique.

It was evident that a photographic system of higher resolving power might be expected to show a greater amount of fine structure in the diagram. This, it was thought, in addition to providing more information, might reduce the ambiguity in the indexing. Moreover, a bundle of fibres could not be maintained easily in perfect parallel alignment, therefore it was evidently desirable to be able to work with a single fibre of the order of 50 μ diameter. In order to use such a small, weakly diffracting specimen without prohibitive exposure times, a small specimen-film distance was necessary. To combine this latter requirement with a system of high resolving power, a well-collimated micro-beam of high intensity was essential.

Micro-Camera and Tube.

A suitable micro-camera was available commercially, manufactured by North American Phillips. A drawing of this camera is shown in Fig.1. It consists essentially of a lead glass capillary collim tor of 10 mm length with either 100 μ or 50 μ bore. One of two alternative flat plate film holders may be set at 15 mm or 10 mm, respectively, from the end of

the collimator. This arrangement was very suitable for work with single fibres which were mounted directly on the flat face of the collimator, and fixed with a cellulose glue.

A micro-beam of high intensity, for the reasons stated above, was the primary consideration in deciding the type of X-ray tube required for this work. With a 100μ diameter collimator inclined at 3° to the target, the projection of the collimator entrance on the target is an ellipse with minor axis 100μ and major axis 2 mm. Therefore a suitable tube would be one with a line focus 0.1 mm by 2 mm and as high a target loading aspossible. Consideration of the tubes available showed that the Ehrenberg-Spear Fine Focus Tube would give the highest intensity beam and least angular divergence with a 100μ collimator, although the focal spot of 50μ diameter was in fact smaller than the optimum size.[7]

Erection of X-ray Tube and Associated Circuits.

Details of the Ehrenberg-Spear fine focus tube have been published[8,9] and it is now manufactured commercially by Hilger and Watts Limited. The tube which was used for most of the work discussed here, was made in the Physics Workshop of Birkbeck College and was one of the few copies to be used outside that department before the complete commercial unit was available. The author would like to acknowledge here the very ready help and advice received from Dr. Ehrenberg and Dr.Spear at the time of the erection of the tube and power supply.

The tube was mounted horizontally on a wooden table; the two windows, one on either side of the target, were 14½" above the table top, which itself was 3' from the floor level. A

Metrovac 02B oil diffusion pump and a Hyvac backing pump provided an adequate vacuum system. The target was cooled by a flow of light transformer oil, which was pumped round a closed series curcuit containing the target and a water cooled Leibig glass condenser.

The high voltage supply was drawn from an ex-Admiralty transformer rated at 45 Kilo-Volts peak output for 230 volts input. This alternating high voltage was rectified by a series chain of six Westinghouse metal rectifiers mounted horizontally at the rear of the tube. The resultant D.C. voltage was smoothed by connecting a condenser, of capacity 0.05 microfarads, between the end of the rectifier chain and earth. This smoothed D.C. voltage was then applied to the anode via a "protective" resistance chain of 10 megohms, so as to prevent any surge of current through the tube if the vacuum should fail. The maximum current that the tube could safely pass, with a 50 μ diameter focal spot, was 500 microamps. Thus the maximum voltage that could be dropped down the protective resistors was only 5 KV.

The control circuit for the target voltage, filament current, oil-circulating pump and diffusion pump was as shown in Fig. 2. With this circuit the usual safety measures were achieved. The main "hold-on" relay, RL_1, ensured that if for any reason the A.C. supply was cut off and then re-established later, none of the tube circuits became alive again until hand operated. The author would like to draw attention to the fact that the voltage supply to a clock, CK, was controlled

by a relay RL3 activated by the rectified A.C. potential drop across a series resistance R_4 in the filament circuit. The filament current was drawn from a 230/12 volt transformer T_2 which depended for its supply on the voltage applied to the primary of the H.T. transformer, T_1, this was cut off if for any reason the tube overloaded. In this way the clock, a 230 volt A.C. synchronous motor, was stopped if the filament burned out or the tube overloaded. This is a feature not generally found in commercial X-ray equipment. Most of the control gear shown in Fig.2. was contained in a chasis 20" by 9" by 9".

With such a small focal spot on the target, and collimators of the same order, it was necessary to avoid any relative motion between the X-ray source and the camera. The micro-cameras were therefore mounted on a heavy piece of iron channelling, 7" wide and 21½" long, which was clamped firmly to the vertical evacuating column of the X-ray tube and set at right angles to the target-filament axis. The top of this "stablizing" girder was arranged to be 8½" below the centres of the two windows on either side of the target.

The X-ray tube and all its associated gear was constructed as a single transportable unit, requiring only a water flow and return, and a 230 volt A.C. supply. The approximate dimensions of this unit were 3' by 3'7" in area and 3' in height to the top of the table; the tube, together with the protective shielding over the target assembly, projected approximately 18" above the table top.

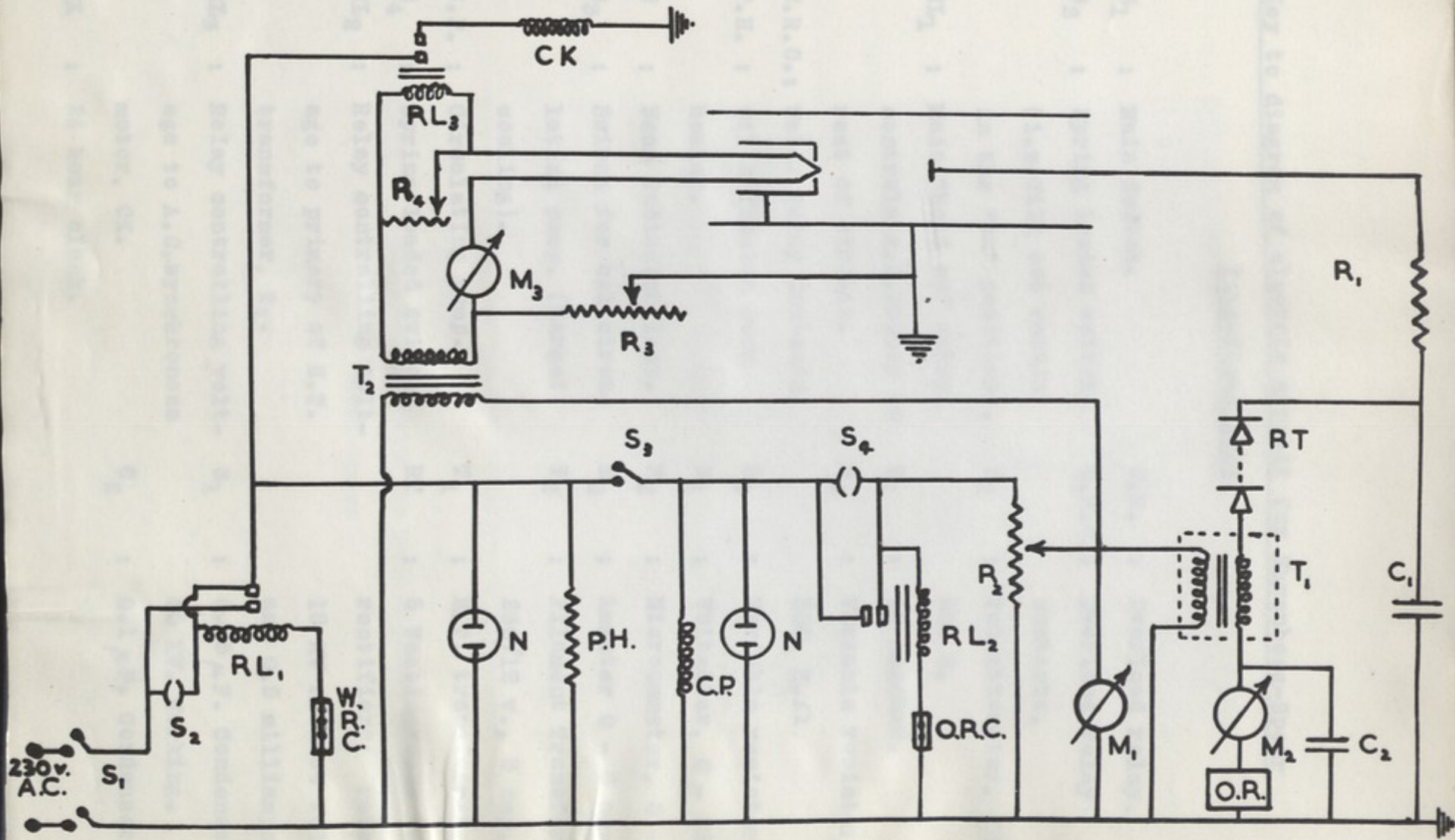


FIG 2: ELECTRIC CIRCUIT FOR EHRENBERG-SPEAR FINE-FOCUS TUBE.

Key to diagram of electric circuit for Ehrenberg-Spear fine-focus tube.

S_1 : Main Switch.

S_2 : Spring loaded switch. (i.e.Will not remain in the "on" position).

RL_1 : Main "hold on" relay; controls A.C.supply to rest of circuit.

W.R.C.: Water Relay Contacts.

P.H. : Oil diffusion pump heater.

N : Neon indicator lamp.

S_3 : Switch for oil circulating pump. (Target cooling).

C.P. : Circulating pump.

S_4 : Spring loaded switch.

RL_2 : Relay controlling voltage to primary of H.T. transformer, T_1.

RL_3 : Relay controlling voltage to A.C.synchronous motor, CK.

CK : 24 hour clock.

O.R. : Overload relay.

O.R.C.: Overload relay contacts.

R_2 : Potentiometer, 1000 Ω. 200 W.

R_1 : 10 megohms.

R_3 : Variable resistor, 300. K.Ω.

R_4 : Variable resistor 7Ω.

M_1 : Voltmeter, 0 - 250 v.

M_2 : Microammeter, 0 - 500 μ.a.

M_3 : Ammeter 0 - 3 amp.

T_2 : Filament Transformer 230/12 v., 3 amp.

T_1 : H.T. transformer.

RT : 6 Westinghouse metal rectifiers; rated at 18 KV reverse voltage and 0.5 milliamps.

C_1 : 0.05 μF. Condenser, 50 KV. Working.

C_2 : 0.1 μF. Condenser.

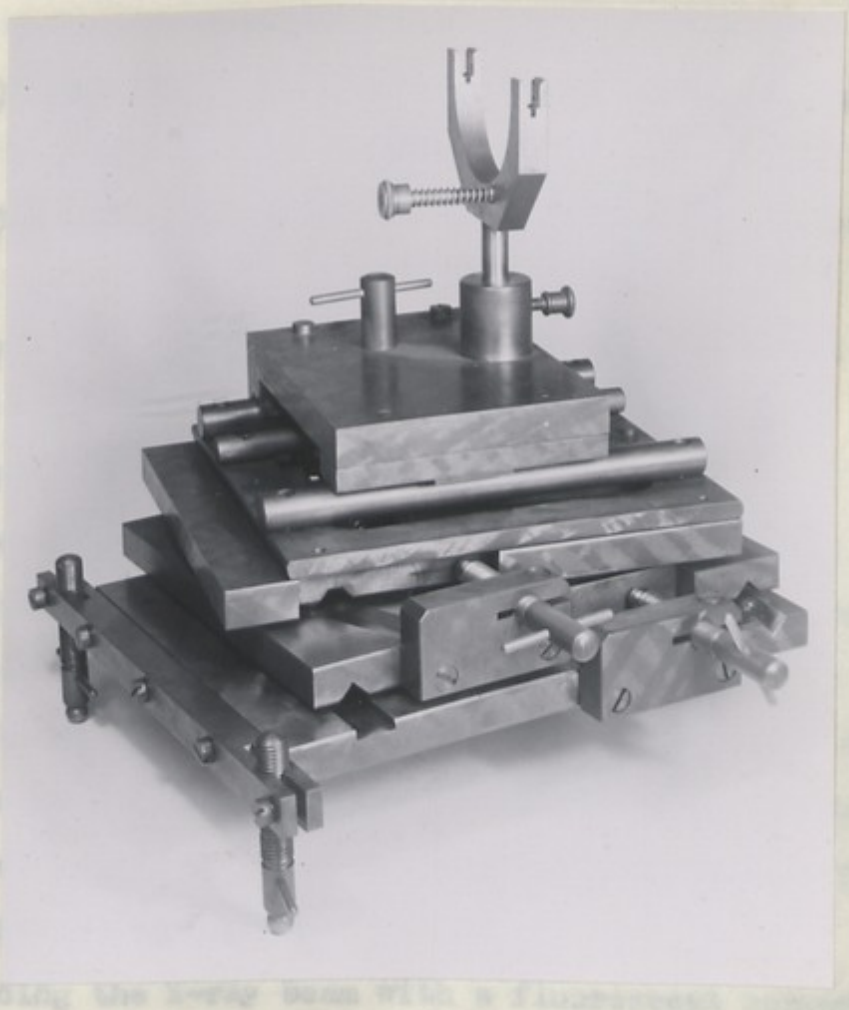


PLATE 7 The stand used with the North-American Phillips Micro-Camera

Camera Stand.

In order to align the N.A. Phillips camera with the X-ray beam it is necessary to mount the camera so that it has the following movements:-

(i) Translation along and at right angles to the X-ray beam.

(ii) A radial movement, preferably about the X-ray source as centre.

(iii) A small angular rotation about a horizontal axis at right angles to the collimator length.

(iv) Usual height adjustments.

The stand shown in Plate 7 satisfied conditions (i), (iii) and (iv) but not (ii). For ease of construction the radial movement was centered on a point mid-way between the front edges of the stand. The rails followed the usual Unicam design and were clamped to the girder bed, attached to the X-ray tube. This stand was designed in conjunction with Mr.L.A. Pitches of the Physics Workshops.

Tracing the X-ray beam with a fluorescent screen was satisfactory with the collimators used (i.e. down to 50μ.) but the use of a Gieger counter detector provided a more rapid method.

Tilting Micro-Camera.

In order to search for further reflections on or near the fibre axis direction, a micro-camera was designed to take X-ray diffraction photographs of a fibre inclined to the X-ray beam. In this way it was possible to explore any particular reciprocal lattice plane right up to the rotation axis (fibre

axis). A general view of this camera and its stand is shown in Plate 8. The specimen carrier and film holder formed one unit in which the film was held in a hemi-cylinder of radius 15 mm. Specimens were mounted so that they lay accurately along the axis of this cylinder; the whole specimen-film unit could be rotated about a vertical axis which intersected the cylinder axis at right angles, so that the unit could be set at any angle to the X-ray beam. A cylindrical film holder was used so that straight layer lines might be obtained whatever the inclination of the X-ray beam to the fibre axis. The collimation was by means of a lead glass capillary tube, made inter-changeable with those of the North American Phillips camera. The collimator was carried by a three plate stand similar to the type used by Hirsch.[10] This stand was set so that the collimator face was 8 mm's in front of the specimen.

(i) Collimator Stand.

The lead glass capillary, approximately 1 cm. long with bore in the range 100 μ to 50 μ, was fixed in a brass mount, Y, which could be inserted in a tube soldered to the plate A shown in Plate 9. Plate A could be moved in either the horizontal or vertical direction by the screws B and C working against the spring D. This assembly was carried by the plate E, which was held, in the vertical plane, by two springs against two locating screws and an adjustable screw F set in a third plate G. This third vertical plate, G, was fixed to the base plate of the camera. In this way the plate A, which carried the collimator, could be given any desired "planar" movement and

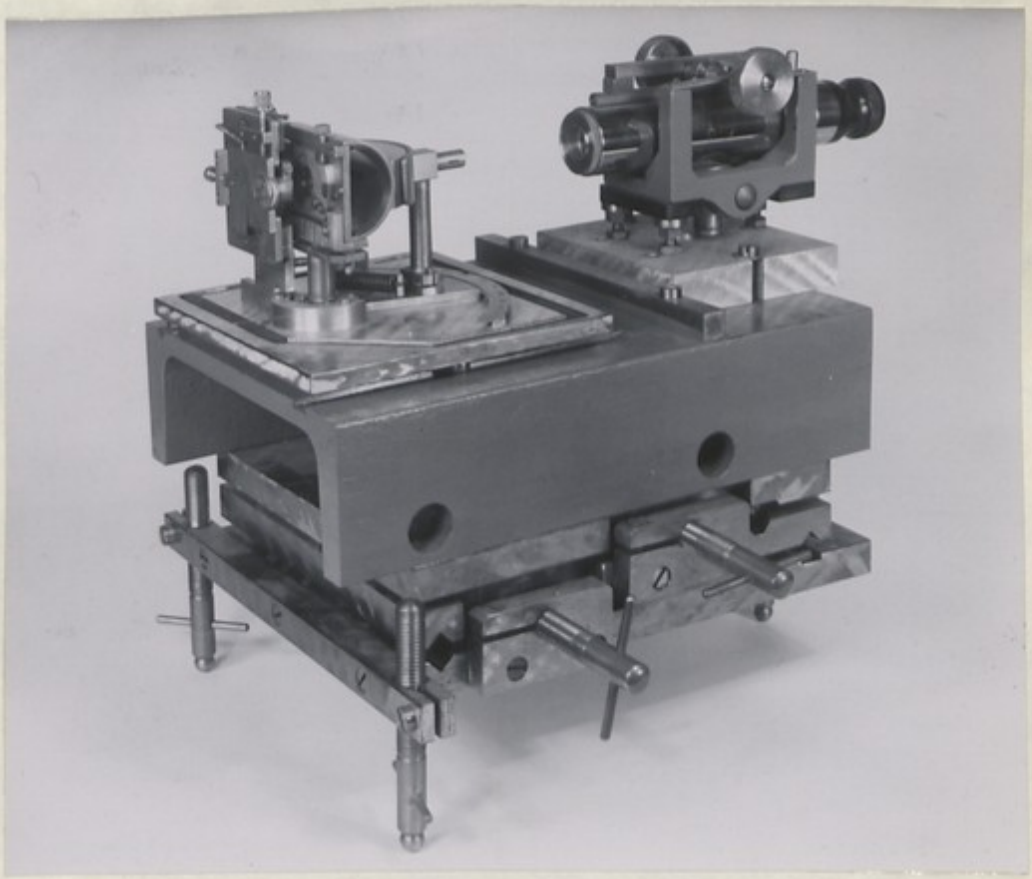


PLATE 8 A general view of the Tilting Micro-Camera and Stand

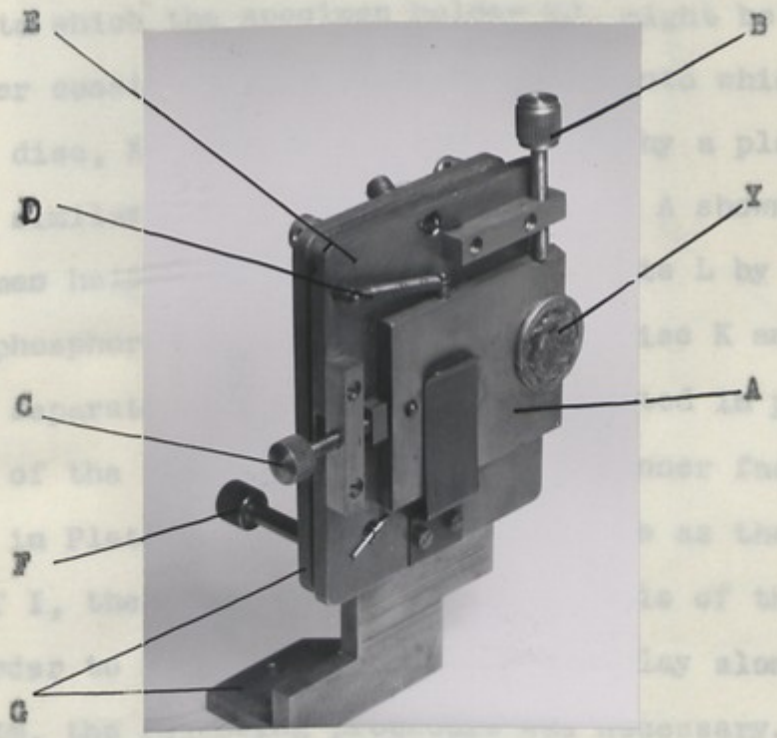


PLATE 9 THE COLLIMATOR STAND

also be tilted to the vertical over a small angular range.

(ii) Specimen - Film Holder Unit.

The flat front face I, Plate 10, of the hemi-cylinder H on which the film was mounted, had its inner surface tangential to the cylinder axis. In the centre of I was a hole into which the specimen holder KJ, might be inserted. This holder consisted of a brass ring, J, into which was screwed a disc, K, the whole being carried by a plate L which had similar planar movements to plate A shown in Plate 9. The specimen holder was retained in the plate L by means of two flat phosphor bronze springs M and N. Disc K and ring J are shown separately in Plate 11; when mounted in plate L the depth of the ring J was such that its inner face, shown uppermost in Plate 11, was in the same plane as the inner surfact of I, therefore it contained the axis of the cylinder.

In order to arrange that the specimen lay along the axis of the film, the following procedure was necessary. The specimen was first mounted taut across the knife edges of the disc K, which was then screwed into the ring J, until the specimen and the inner face of J appeared to be in the same plane as judged by viewing them under a microscope. Then on replacing the specimen holder KJ, in the plate L the specimen was in the same vertical plane as the axis of the cylinder containing the film. It was then required to rotate the ring, in plate L, until the fibre was seen to lie along the vertical axis of rotation of the whole unit, when viewed through the low power microscope mounted on the camera stand

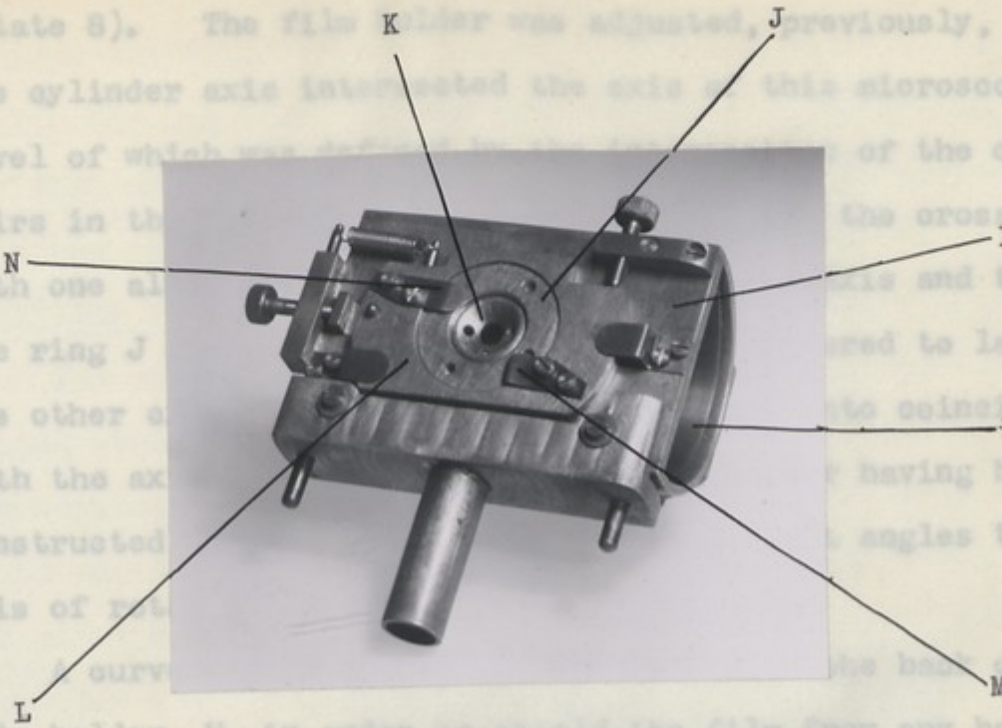


PLATE 10 THE SPECIMEN-FILM HOLDER UNIT

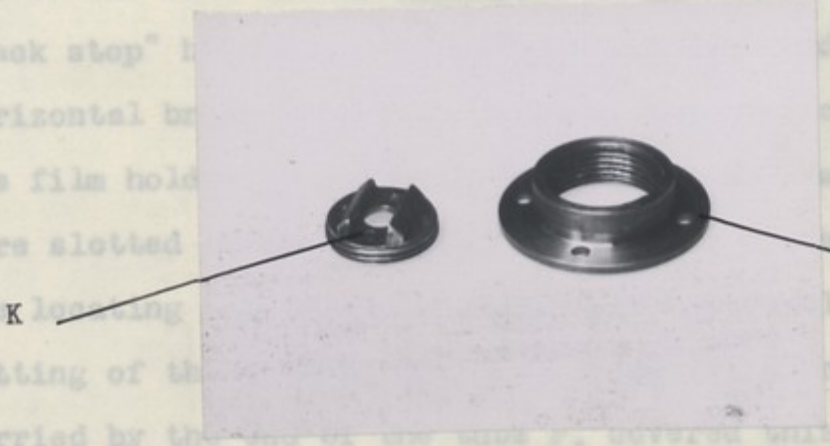


PLATE 11 THE SPECIMEN HOLDER

(Plate 8). The film holder was adjusted, previously, so that its cylinder axis intersected the axis of this microscope, the level of which was defined by the intersection of the cross-hairs in the eyepiece. Therefore by setting the cross-hairs with one along the direction of the rotation axis and turning the ring J through 90° so that the fibre appeared to lay along the other cross-hair, the fibre was brought into coincidence with the axis of the cylinder; the film holder having been constructed to have its cylinder axis at right angles to its axis of rotation.

A curved brass block, O, was screwed to the back of the film holder, H, in order to shield the film from any back scattered X-radiation. The direct beam was allowed to pass out of the film holder through a hole in the film. This "back stop" hole was located by a pin inserted through a horizontal brass tube P, mounted at the correct height behind the film holder. The backing block O and the film holder were slotted on one side, as is shown in plate 12, to enable the locating pin to position the film correctly at any angular setting of the unit to the direct beam. A brass shield, carried by the end of the tube P, covered this slot in the film holder. The collimator - specimen - back-stop axis was defined by the axis of the "back-stop" tube P, as is shown in Plate 13. The axis of the low power microscope (50 mm. objective), shown in Plate 8, was made to coincide with this "back-stop" axis by focusing on a small pin hole, 100μ diameter, in a single ended brass tube which could be set at

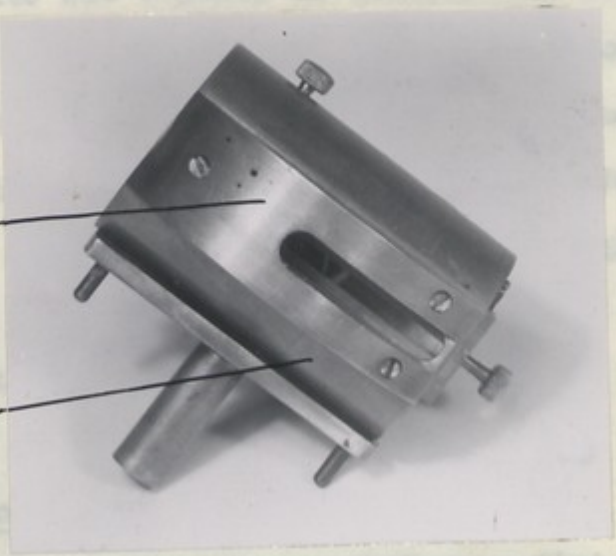


PLATE 12 A rear view of the Specimen-Film Holder Unit, H, showing the backing block O and the slot in H and O to permit the escape of the direct beam at any angular setting of the unit.

various positions in the tube P. The alignment of the collimator and specimen was then achieved by adjusting their position until an image of each in turn could be focussed onto the cross-hairs of the microscope.

(iii) The Specimen-Film Holder Mounting.

The specimen-film holder was carried by a vertical T-shaped stand R, attached directly to a sector plate S, in the horizontal plane, as is shown in Plate 14. This sector plate was held by a strong spring, T, against a lever, U, in the plane of the plate but capable of independent rotation about the same axis as the stand R. Thus the sector plate could be moved round by the lever and the angular rotation from the zero position, i.e. X-ray beam normal to the cylinder axis, read off on a scale, V, marked directly in degrees. A fine adjustment was available by clamping the lever and rotating the sector plate by working a screw, W, in the lever arm against the tension of the spring T.

(iv) Camera Plate and Stand.

The units of the camera were carried on a flat rectangular base plate which had a channel cut around the edge, as may be seen from Plates 13 and 14. In this channel, which was rubber lined, a heavy brass cover could be seated, so that the camera might be evacuated. Alternatively the air could be replaced by hydrogen at the desired relative humidity, when the water content of the specimen was critical. One corner of the base plate was cut away to allow the camera to approach more closely to the window of the tube.

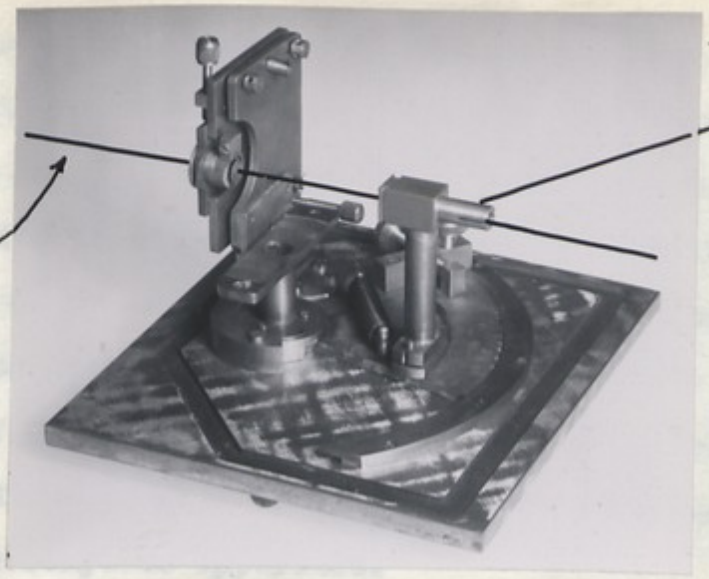


PLATE 13 Base plate of camera showing collimator stand and backstop tube, P. The axis of P defines the alignment axis of the camera

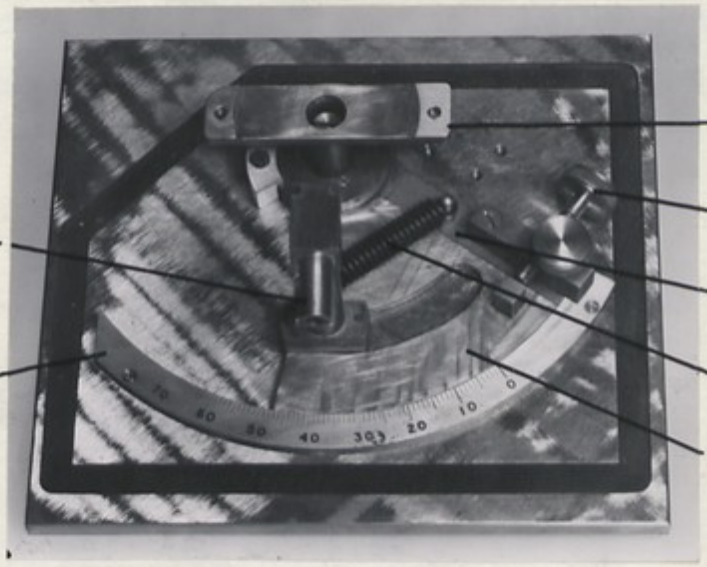


PLATE 14 Base plate of camera showing specimen-film holder mounting R, backstop tube, P, and scale V. The collimator stand shown in Fig.13 has been removed.

The camera was mounted on a similar stand to that used with the North American Phillips camera. In order that the camera might be aligned at the X-ray tube, removed, loaded and then replaced in correct alignment, the three half-ball bearing feet of the base plate were located on the supporting stand by means of the usual three point kinamatic seating, consisting of a groove, flat and a hole. The stand was designed by Mr. L.A. Pitches of the Physics Workshops who also made the complete apparatus.

REFERENCES IN CHAPTER III

1. Signer, R. and Schwander, H.(1949). Helvet.Chi.Acta.32. 854.
2. Signer, R. and Schwander, H.(1950). Helvet.Chim.Acta.33.1522
3. Wilkins, M.H.F. (1949): Private Communication.
4. Wilkins, M.H.F. Gosling, R.G. and Seeds, W.E. (1951): Nature. 167. 759.
5. Astbury, W.T., (1947). S.E.B. Symp. 1. 66.
6. Bunn, C.W. (1945): Chemical Crystallography. ChapterIXI. Published by Clarendon Press.
7. Huxley, H.E. (1953) Acta Cryst. 6. 457.
8. Ehrenberg, W., and Spear, W.E. (1951): Proc. Phys. Soc. B. 64. 67.
9. Spear, W.E. (1951): Ph.D. Thesis. London.
10. Hirsch, P.B. and Kellar, J.N. (1951): Proc. Phys. Soc. B.64. 369.

CHAPTER IV

THE X-RAY DIFFRACTION PATTERNS FROM NaDNA FIBRES, THEIR DEPENDANCE ON FIBRE WATER CONTENT AND THE QUALITATIVE INTERPRETATION OF THE EXPERIMENTAL DATA.

The majority of the X-ray diffraction photographs discussed here were taken with the N.A. Phillips Micro-Camera, and the Ehrenberg-Spear tube described in the previous chapter. For most of the work the 100 μ collimator and 15 mm specimen-film distance were used, together with nickel-filtered copper radiation. The Ehrenberg-Spear tube gave a focal spot of about 50 μ in diameter and was run at 35-40 KV. and 0.4 ma. The specimens were generally single fibres of 30 - 40 μ in diameter, requiring exposure times of about 50 - 100 hours. For fibres of diameter about 100 μ exposure times were 20 hours or less.

Preliminary investigations, described in Chapter III, had shown that it was important to eliminate the air scattering and also that the structure was very sensitive to water content. Therefore, throughout each exposure, a steady stream of hydrogen at constant humidity was passed through the camera. This was achieved by bubbling the hydrogen through a saturated solution of a suitable inorganic salt. A list of the salts used, and their relative humidities, is given in Table 1. A little of the same saturated solution was also placed in a small container inside the camera. Each specimen was left for at least one hour at a constant humidity within the camera, before starting an exposure.

Phase Change with Water Content:

Structure A and Structure B

It had been observed (Chapter III Plates 4, 5 and 6) that a high ambient humidity was required before the NaDNA fibres gave a diffraction pattern indicative of a fairly high degree

TABLE I

This table shows the relative humidity, at a given temperature and in an enclosed spaced, above the saturated salt solutions used to obtain constant humidity around the specimen during an exposure.

Salt	t°c.	R.H. %
NaCl	16.4	30.5
$CaCl_2$. $6H_2O$.	18.5	35
$Ca(NO_3)_2$. $4H_2O$.	24.5	51
$Ca(NO_3)_2$. $4H_2O$.	18.5	54
$NaNO_2$	20	66
NaCl O_3	20	75
Na_2 CO_3 $10H_2O$.	24.5	87
Na_2 CO_3 $10H_2O$.	18.5	92

of crystallinity. Attempts to achieve a further increase in crystallinity by working at still higher R.H. resulted in the observation that, at very high humidities, the NaDNA molecular structure under-goes a well defined change, leading to a new type of fibre-diagram. The two types of diagram are shown in Plates 1 and 2. The molecular structures giving rise to these two diagrams will be referred to as Structure A and B.

On passing from Structure A to Structure B the layer-line spacing increases from 28A° to 33A°, and this is accompanied by a similar 20 - 30% increase in the macroscopic length of the fibre. This change from A to B is, in general, readily reversible, depending chiefly on the water content of the fibre but also to a certain extent on the past history of the specimen. A certain amount of hysteresis was observed in the quantity of water taken up by bulk specimens of NaDNA and also in the structural change of fine fibres as indicated by their X-ray patterns. Thus a diagram similar to Plate 1 may be obtained frequently at 92% R.H. if the specimen has been strongly dried previously over P_2O_5.

The highly order Structure A had been obtained in non-orientated form, by Riley and Oster.[1] The spacings reported by them for the "crystalline" structure, in the Signer and Schwander preparation, agree to within 0.1A° with those of strong single reflections and also with the mean values of doublets and triplets in the fibre diagram of Structure A (Plate 1). Riley and Oster also observed that the transition from this state to the "wet gel" was reversible. However, no

PLATE 1

Three fibres of NaDNA; diameters 18-30μ. 15 mm. specimen to film distance. Exposure 116 hours. R.H. 75%

STRUCTURE A

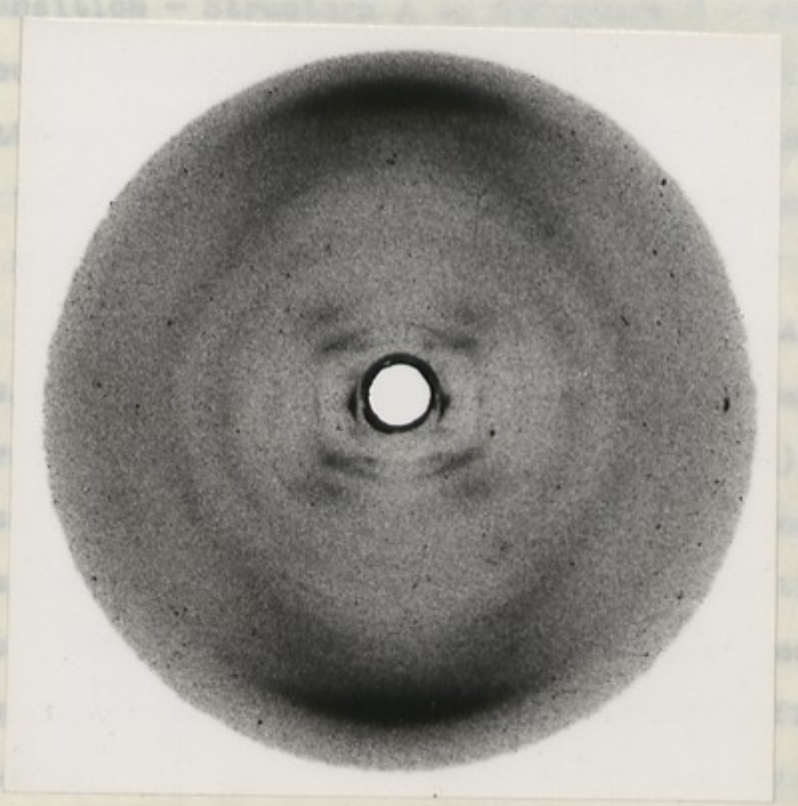


PLATE 2

Several fine fibres.

Specimen to film distance 15 mm.

Exposure 6 hours.

R.H. 92%

This specimen had previously given patterns similar to Plate 1.

STRUCTURE B

measurements in the high-angle region with "wet gel" or "moist crystalline" specimens were made, therefore it was not observed that this reversible change was an intramolecular transition between two distinct states. They interpreted their data, from observation of the changes in the diffraction pattern at low angles, in terms of inter-molecular changes due to miscelle formation, or as an alternative, a continuous intra-molecular change. This latter interpretation required the molecules to be regularly coiled or folded in such a way that their diameters, as well as the repeat distance along their axes, expanded on increasing the hydration of the specimen. It would seem from the work presented here that the intramolecular change is not continuous, but that a discrete transition - Structure A to Structure B - occurs in NaDNA specimens containing above approximately 40% by weight of water. Although the axial repeat period of B is greater than that of A, the diameter of the molecules probably decreases at the A to B transition. (Chapters V and VI).

Using some samples of calf-thymus NaDNA, although it was possible to obtain well-orientated fibres (as shown by optical birefringence and X-ray diffraction pattern), it was not found possible to obtain Structure A. Fibres from such samples gave diagrams indicating Structure B at all relative humidities above about 60%. The pattern gradually became less distinct on drying comparable to the behaviour of Structure A fibres. The failure to obtain Structure A in these specimens may have been due to degradation during extraction of the DNA, though this seems

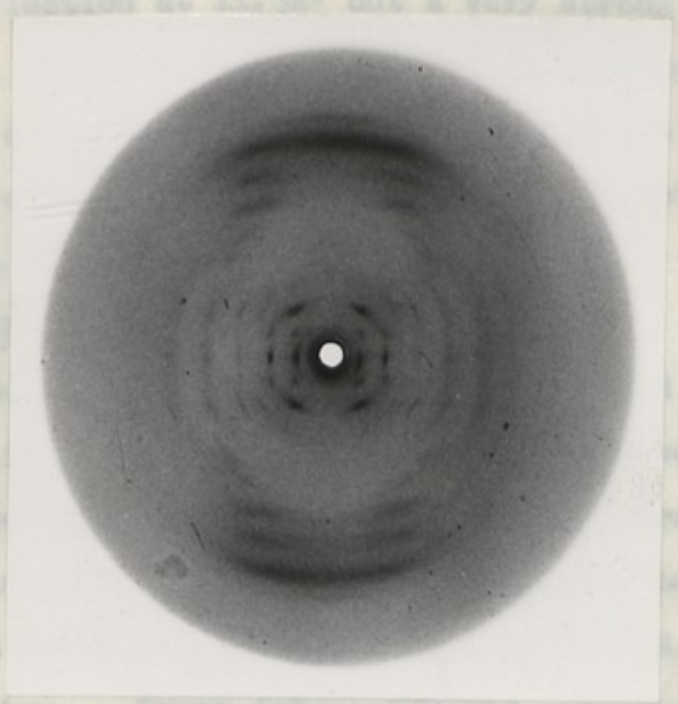
unlikely, or to the presence of some other material, perhaps traces of protein, which could cause a certain degree of cross-linking so preventing the contraction of the molecule into Structure A.

Reversibility of Phase Change A to B.

For specimens in which Structure A could be obtained, when R.H. 75% was approached from the dry side, only, Structure A was formed. When, with the same specimen, R.H. 92% was approached from the wet side the resulting diffraction pattern indicated a complete structural change, showing only Structure B. At intermediate humidities mixtures of A and B were obtained.

Mixture Pattern A plus B.

An example of this "~~mixing~~ mixture pattern" is shown in Plate 3. The published fibre diagrams of Astbury et al [2.3.] were all of Structure A plus a small amount of Structure B. One of the differences between the two patterns is the complete absence of a strong meridional arc in the pattern of Structure A, whilst the meridional arc at 3.3A° is one of the most striking features of the diagram of Structure B. A mixture of A and B therefore suggests that an apparent layer line period of 27.5 - 28A° (due to A), is to be associated with a repeat spacing of 3.3A° (due to B), both along the fibre axis. It was this interpretation of the mixture as a single phase that caused Astbury to suggest that the arrangement in space of the nucleotides was repeated along the fibre axis every eight nucleotides, taking the effective thickness of a nucleotide as about 3.4A°.



MIXTURE PATTERN

PLATE 3 Single fibre, diameter 50μ.

Specimen-film distance 15 mm; exposure 30 hours.

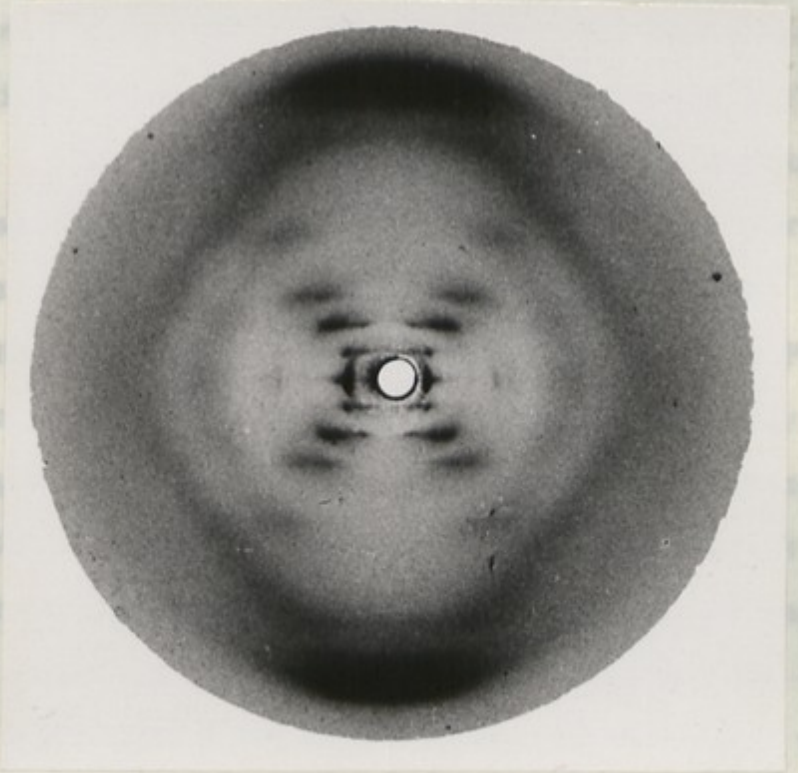
R.H. 75%

The presence of a small amount of Structure B in a pattern of A is most readily detected by the effect structure B has on the innermost equatorial reflection, circa 20A°. In the Structure A pattern this reflection is very weak and the second equatorial reflection, 11.3A°, is very strong (Plate 1). The Structure B pattern however, shows no equatorial reflection at 11.3A° but a very strong reflection at about 20A°, or higher, depending on the hydration. Thus a combination of A and B often produces a pattern in which the two innermost equatorial reflections have high intensities as is shown in Plate 3.

Irreversible Change A to B.

It was found that sometimes a fibre that had been in use in the camera for some days (in some cases even weeks), and which had passed through the reversible A to B transformation a number of times, suddenly passed irreversibly to Structure B. The photograph shown in Plate 4 was obtained at R.H. 75% from such a fibre. These fibres behaved subsequently like those mentioned above from which Structure A could not be obtained.

The difference between Plates 2 and 4 was largely due to the difficulty of photographing a very wet fibre in a highly orientated state. A fibre under-going the A to B transformation suffers a length increase of about 25%. Therefore in order to follow the process with the same specimen it was necessary to attach the fibre to the collimator at points very close to the capillary bore, so that the fibre could not move out of the direct beam. Even so, some buckling of the



STRUCTURE B

PLATE 4 Single fibre, diameter 50μ

Specimen to film distance 15 mm; Exposure 62 hours

R.H. 75%.

Result of irreversible A → B change.

fibre, did occur in the wet state with consequent deterioration of the diffraction pattern recorded. No such length change and hence no such difficulty was encountered with fibres which did not give Structure A.

Structure Below 75% R.H.

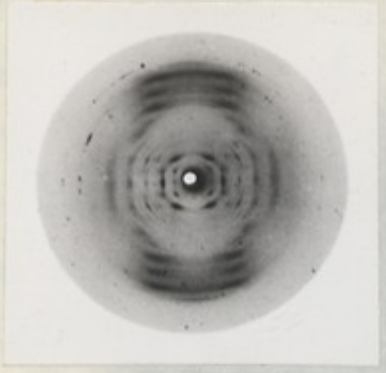
If with a specimen showing Structure A the humidity in the camera was appreciably reduced from 75%, the intensity of the reflections decreased relative to that of the diffuse background. This effect increased progressively with decreasing humidity and was accompanied by a slight shrinkage of the structure, mainly in a lateral direction. This can be seen from the four diffraction photographs shown in Plate 5. After prolonged drying at room temperature, (c), only a broad diffuse ring was observed. If the specimen was dried at 80°C, (d), the diffuse ring at about 4 - 3.4A° remained and in addition strong diffuse scattering appeared at low angles.

As was discussed in Chapter III the change from the crystalline Structure A to this disordered state was wholly reversible.

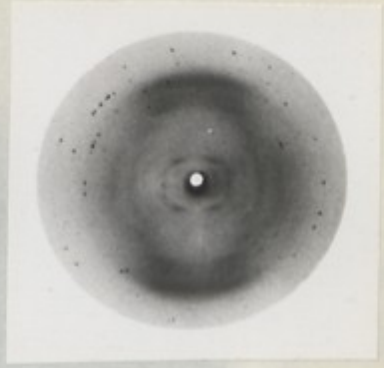
Double Orientation in Structure A.

By chance, it was found that one fibre of 40μ diameter gave a Structure A pattern showing strong double orientation, Plate 6. That is, the crystallites were not in random orientation about the fibre axis, therefore the resulting X-ray pattern was something intermediate between a rotation and an oscillation photograph. It may be seen from Plate 6 that many pairs of equivalent reflections in adjacent quadrants

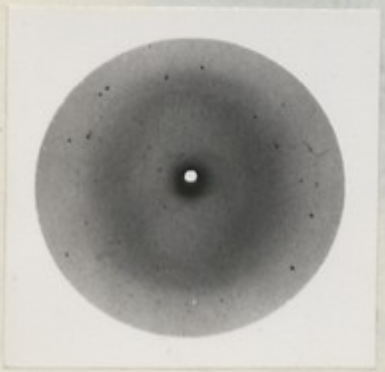
PLATE 5: STRUCTURE A BELOW 75% R.H.



(a) Single fibre. Diameter 120μ. Exposure 22 hours. Relative Humidity 56%

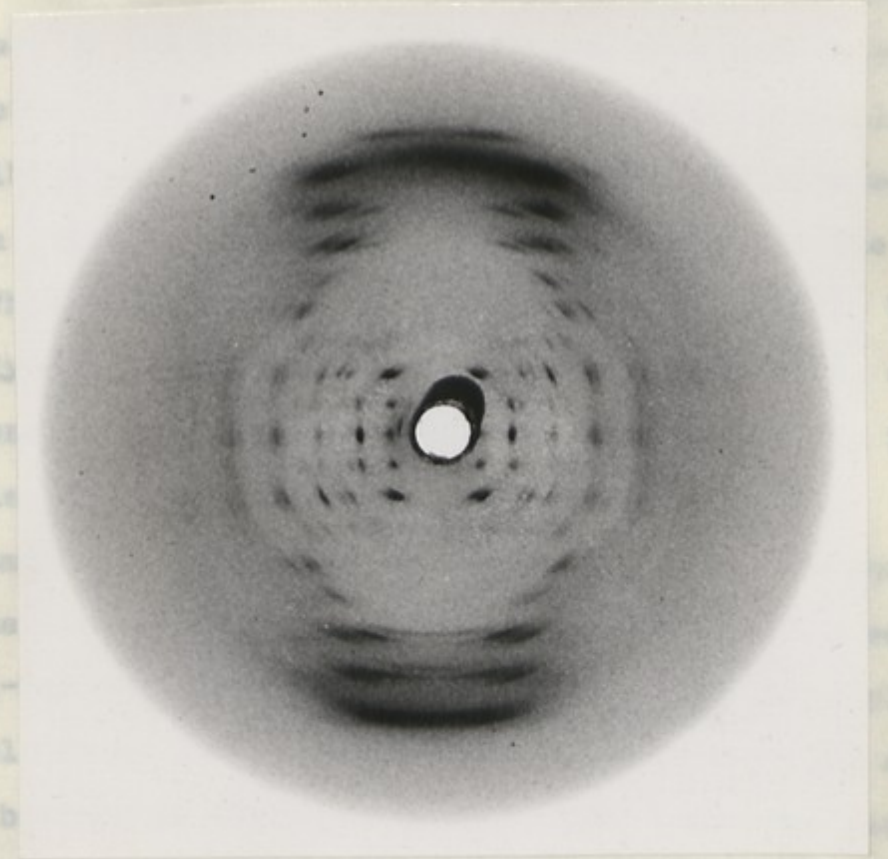


(b) Specimen as for (a) Exposure 19 hours Relative Humidity 40%



(c) Specimen as for (a) Exposure 24 hours. Dried over P_2O_5 at room temperature.

(d) Circa 35 fibres. Diameter 10-30μ. Unicam camera. Specimen dried over P_2O_5 at room temperature and then heated in camera to about 80°C prior to exposure.



DOUBLE ORIENTATION OF STRUCTURE A

PLATE 6 Single fibre, about 40μ diameter.

Specimen-film distance 15 mm; R.H. 75%

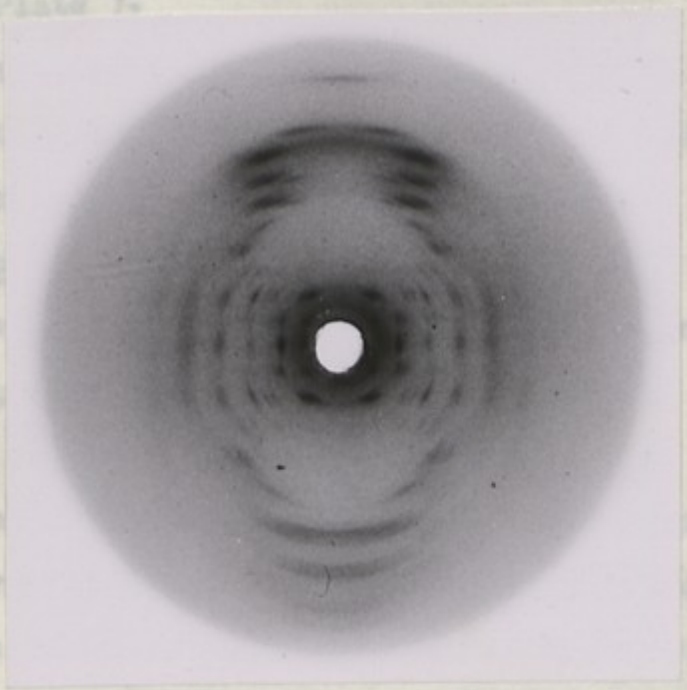
Exposure about 100 hours

have marked different intensities.

It was thought that this double orientation effect was probably due to either, a mechanical accident to the fibre, or a preferential orientation of the crystallites which lie near the surface of the fibre. Two series of attempts to reproduce the effect were therefore made, both being unsuccessful. First, single fibres were placed on a glass slide and flattened by rolling a glass rod along them, in the direction of the fibre axis. When this was done at room humidity the fibres became positively birefringent, as first observed by Wilkins[4], and gave a disordered diffuse diffraction pattern. Therefore, fibres were rolled while at a relative humidity near to 75%. The fibres remained negatively birefringent but they subsequently showed the less ordered Structure B pattern. In the second series, X-ray photographs were taken of single fibres placed on the collimator so that only a small region near the surface of the fibre was in the X-ray base. This technique gave several well-oriented photographs of Structure A, but none with any traces of double orientation.

Meridonal Reflections in Structure A.

The tilting micro-camera described in Chapter III was built primarily to search for extra reflections close to the fibre axis in the A structure. To obtain a reflection with reciprocal space co-ordinates $\xi = 0$ and $\zeta = \zeta_1$, the rotation axis (i.e. fibre axis) should be inclined at an angle γ to the direct beam, where $\sin\gamma = \zeta_{1/2}$. Therefore a fibre giving a good A pattern was selected and a series of exposures were taken, with values of γ ranging from $4^{\circ}30'$ to $19^{\circ}\ 30'$. The



PATTERN SHOWING 11th LAYER LINE MERIDIONAL REFLECTION

PLATE 7 Single fibre, diameter 80μ.

Specimen-film distance 15 mm; Exposure 119 hours

R.H. 75%

This film was taken with the tilting micro-camera described in Chapter III. In this instance the specimen and the cylindrical film holder were inclined at $14\frac{1}{2}°$ to the X-ray beam.

only extra reflection observed with this technique was the meridional arc on the 11th layer line. In none of the many exposures showing this reflection was their any indication that the arc might be resolved into two discrete reflections. The pattern obtained for $\gamma = 14^{\circ}\ 30$, corresponding to γ_9, is shown in Plate 7.

The extra resolution produced on the higher layer lines by these tilting experiments was useful in indexing the patterns obtained with the Phillips camera.

Disintegration of NaDNA fibres by X-rays.

On a few occations a fibre which had been in use for some-time, showing a well orientated structure, would give a powder diagram of a few sharp rings similar to the type of pattern to be expected from small crystals of an inorganic salt. This powder pattern was always superimposed on a diffuse background, an example is shown in Plate 8. With a fibre that was com-pletely in the X-ray beam, the fibre either disintegrated or a sharp neck was seen where it crossed the collimator capillary; with a fibre larger than the beam the X-rays appeared to drill a hole through the fibre, as is shown in Plate 9.

It would seem, therefore, that irradiation with X-rays can cause a collapse of the NaDNA structure. Whether the effect is cumulative over successive exposures, or is a spontaneous chain reaction through the irradiated area, or a combination of both effects, is not known. However, the few fibres in which the effect has been observed showed no succes-sive deterioration of diffraction pattern prior to their collapse.

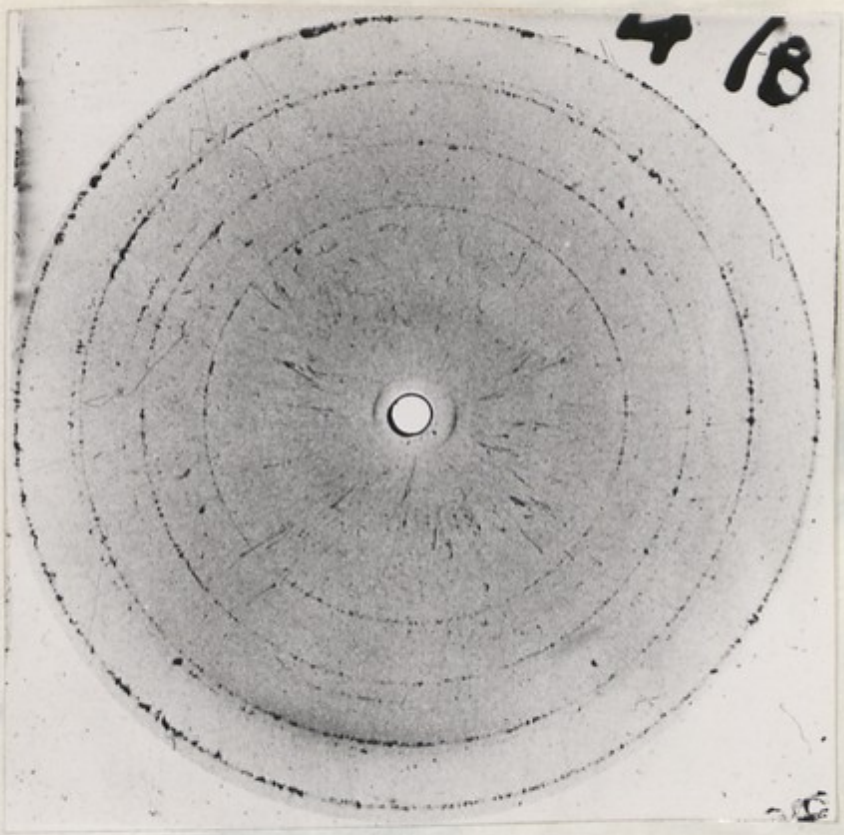


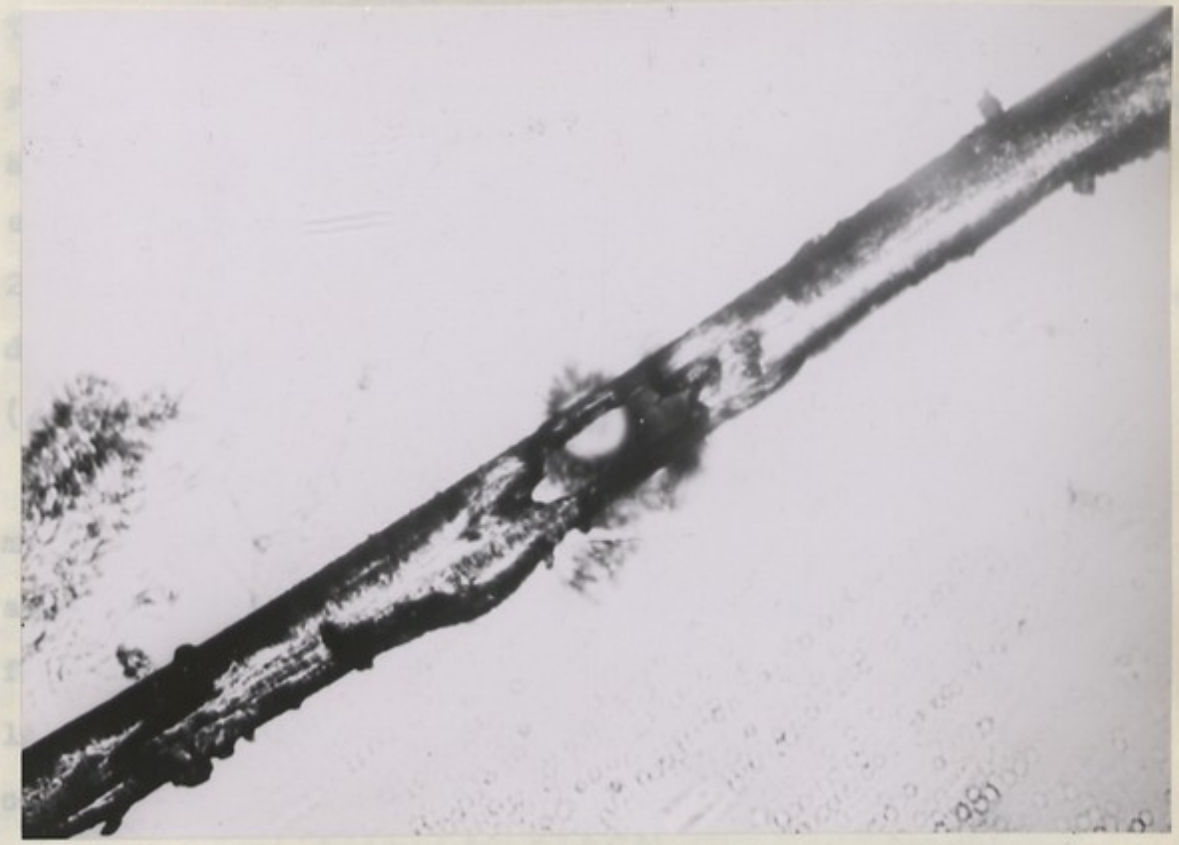
PLATE 8 A powder photograph obtained from a "collapsed" fibre of NaDNA.

The spacings corresponding to the diffraction rings shown in this photograph are listed below, together with an indication of their intensity.

3.15A°. (M)	2.85A°. (V.W.)	2.55A°.(M)
2.45A°. (W)	2.25A°. (S)	2.20A°.(W)
2.00A°. (S).		

Intensity Scale S : Strong W : Weak

M : Medium VW : Very Weak



DISINTEGRATION OF NaDNA FIBRE BY X-RAYS

PLATE 9 A photograph of a NaDNA fibre mounted directly on the collimator face. The irregular hole in the fibre, above the 50μ capillary bore of the collimator, occurred during the exposure.

Water Content and Density Determinations.

For a quantitative interpretation of the X-ray data it is important to know both the water content and the density of the crystallites in these fibre specimens. Although it was possible to measure the water up-take of fibres, it was not found possible to measure their densities at various humidities. Therefore it was necessary to use "bulk specimens". Homogeneous pellets of dry NaDNA were prepared by allowing pieces of clear gel to dry slowly in air, taking care to remove any trapped air. These air-dried pellets were then dried over P_2O_5 at 25°C for several weeks, or in some cases, months. The average dry weight of the pellets used was 5 milligrams.

(i) Dry Density.

All density measurements were made by a "float-sink" method, in which the density of the liquid was varied by altering the temperature. The dry density was measured by flotation in carbon tetrachloride which has a density of 1.632 g/cc at 0°C. Each pellet was placed in a test tube, containing some carbon tetrachloride and the temperature allowed to rise slowly, from about - 10°C, until the pellet just sank; the temperature of the liquid at this instant was noted. Care was taken to keep the pellet a little beneath the surface whilst the temperature of the liquid was below the dew point. In this way a density range of about 1.65 to 1.54 g/cc could be covered, corresponding to temperatures of - 10°C to 30°C.

The density of dry NaDNA, as determined by the above method, was 1.625 $\pm$ 0.002 g/cc at 4 $\pm$ 1°C. This finding is

in good agreement with Astbury's value[3] of 1.63 g/cc. It was, moreover, the density of most of the pellets made; any other value was lower than 1.625 and was probably due to trapped air or other contaminate.

(ii) Density and Water Content at 75% R.H.

In all cases the dry density was determined before any water up-take measurements were made. Only pellets having a dry density of 1.625 ± .002 g/cc were used. These pellets were maintained at the required relative humidity until equilibrium was reached, as indicated by no further weight change. The pellets were weighed correct to $\pm 10^{-5}$ g using a Stanton Micro-balance. The weighings were carried out as rapidly as possible because of the ready water exchange between NaDNA and the surrounding atmosphere. With dry pellets it was possible to observe a change of $+ 10^{-4}$ g in 2 or 3 minutes, if the pellets were allowed to remain on the balance pan.

Several pellets were allowed to equilibrate at 75% R.H. for about six weeks. The water up-take of these pellets was found to be 37% of the dry weight with a variation of only ± 1%. The densities were determined, by flotation in chloroform, to be 1.471 ± .002 g/cc at 30°C ± 1°C. By using chloroform over the temperature range from - 10°C to + 30°C densities from 1.54 to 1.47 g/cc could be measured. This range overlapped with that of the carbon tetrachloride density between the same temperature limits. Thus, by using carbon tetrachloride and chloroform, a continuous range of densities

from 1.65 to 1.47 could be covered with an accuracy of ± .002 g/cc, corresponding to an error of ± 1°C in the "just sinks" temperature.

The diffraction pattern of several of these pellets was obtained and found to correspond, almost entirely, to Structure A. This was indicated by the absence of strong reflections corresponding to spacings greater than 12A°.

It was found that the water-content of the pellets was always lower than that of fibres at the same humidity. Fibres at 75% R.H. were found to take up about 42 - 45% by weight of water.

(iii) Density and Water Content at 92% R.H.

The pellets used for determinations at 75% R.H. were maintained at a relative humidity of 92% for more than nine weeks before a density determination was made. Toward the end of that time the weighings were constant to within ± 5 10^{-5}g. The water up-take was found to be 60 ± 2% for all the pellets used.

For the determination of density by flotation method, a liquid is required which will not absorb water. A liquid with a density in the required range and an accurately known temperature coefficient of expansion is p-bromotoluene, which has a melting point of 28°C. Using this liquid and the float sink method, the density of the pellets was found to be in the range 1.380 ± .002 g/cc corresponding to a temperature range of 41°C ± 2°C. These pellets at 92% were used as specimens in the micro-camera and gave powder diagrams showing spacings at about 22A°. This indicated that the A to B transformation had taken place in these bulk specimens and that the density of

1.380 g/cc was related to Structure B.

Relation of Bulk Specimen to Crystallites.

The equilibrium water content of NaDNA pellets at 75% R.H. was found to be fairly constant for a number of different specimens. However, it is likely that the water up-take of the crystallites was different from that of the amorphous regions. It may be that the strong orientating action of fibre drawing increased crystallite formation compared with the slow drying of a gel, so that the Crystallite/Amorphous weight ratio in a fibre was higher than that of a dried pellet of gel. If this were true, since the water content of fibres was always found to be higher than that of pellets at the same humidity, it may be concluded that the water content of the crystallities was greater than that of the inter-crystallite regions. This is in accord with the work of Wilkins[5], who has concluded that the water up-take of specimens with a low Crystalline/Amorphous weight ratio is appreciably smaller than that of specimens with a high value of this ratio. An alternative explanation is that there may have been excess water in layers on the surface of the crystallites. However, if the former explanation is accepted then the water content of bulk pellets must be regarded as a minimum value for the crystallites. The density value of 1.471 g/cc found for pellets at 75% R.H. is accordingly a maximum value for the density of crystallites, at that humidity.

Qualitative Interpretation of Experimental Data.

One of the first structural suggestions that may be made, from the form of the diffraction patterns presented above, is that Plate 4 (Structure B Pattern) is strongly characteristic of the type of diagram shown by Cochran, Crick and Vand,[6] (simultaneously and independently derived by Stokes[7]), to result from a helical structure. The quantitative interpretation of this pattern on a helical basis is discussed in Chapter V.

With regard to both Structures A and B; the most polar parts of the NaDNA molecule are the phosphate groups, it is, therefore, with these groups that the water molecules are most likely to be associated. It may be noted, for example, that the compounds $(C_2H_5O)_2 PO_2 Na$ and $(C_3H_7O)_2 PO_2 Na$, in which the bonding of the phosphorus is similar to that in NaDNA, are highly hygroscopic.[8] Because of the polar nature of these phosphate groups, it is likely that they associate with each other in the crystallite structure, in a manner similar to carboxyl groups in crystals of organic acids and their salts. That is, intermolecular links of the type shown in Fig.1. may be expected.

(i) Inter-molecular Phosphate - Phosphate Bonds.

The extreme rapidity with which water is taken up or lost by NaDNA fibres when the humidity of the surrounding atmosphere is changed, the large quantity of water which can be absorbed and the continuity of the process, which leads ultimately to gel formation and solution, suggest that the phosphate-phosphate intermolecular links postulated above are in a relatively accessible part of the structure. When the water

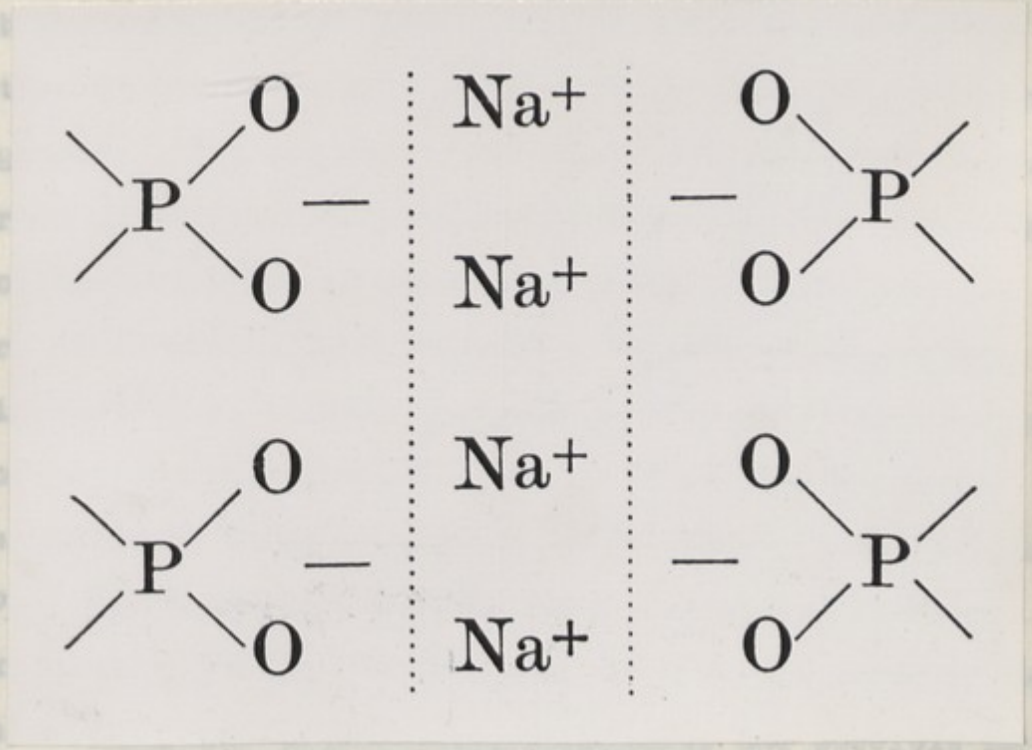


FIG. 1 Suggested inter-molecular bonding between phosphate groups of neighbouring polynucleotide chains

content of those NaDNA fibres that show Structure A is sufficiently increased, Structure B always results. It is clear from the X-ray diagrams that Structure B has a lower degree of order than Structure A. This suggests that the additional water has in some way weakened the directional property of these phosphate-phosphate bonds. This could occur by water molecules associating themselves directly with the sodium ions and forming a new intermediate in the link, which would make possible some degree of intermolecular displacement in the direction of the fibre axis as well as an intra-molecular rearrangement. The suggestion of water as an intermediate is supported by the photograph shown in Plate 10 (a). Here the equatorial reflection at about 22A°, which is a prominent feature of Structure B, appears as a well-resolved doublet. This is shown more clearly in Plate 10 (b), which is an enlargement of the centre region of Plate 10 (a). The spacings corresponding to these equatorial reflections are 21.2A° and 23.7A°. The difference between these spacings, 2.5A°, is approximately equal to the thickness of a single layer of water molecules and seems to indicate, therefore, the possible co-existence in the structure of at least two different hydration states.

(ii) NaDNA Structure in Solution - Structure B.

The water content of fibres showing Structure B may vary within wide limits so that it would seem that Structure B can co-exist with a large quantity of inter-micellar water. Increasing the water-content beyond the lower limit necessary for the A to B transformation increases lateral swelling of the

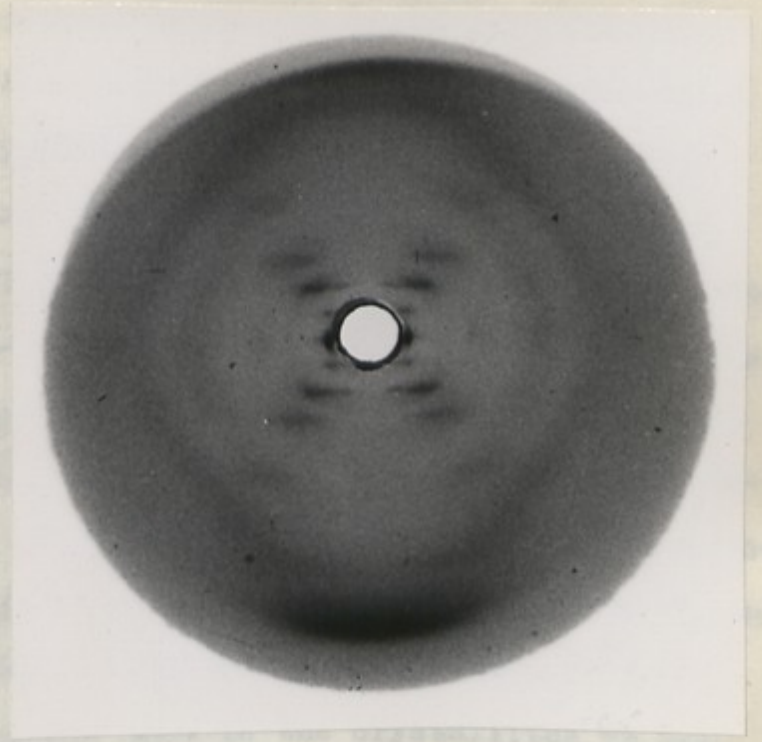


PLATE 10 (a)

Single fibre,

50μ diameter.

Exposure 62 hours.

R.H. 75%

Result of irreversible

A → B change.

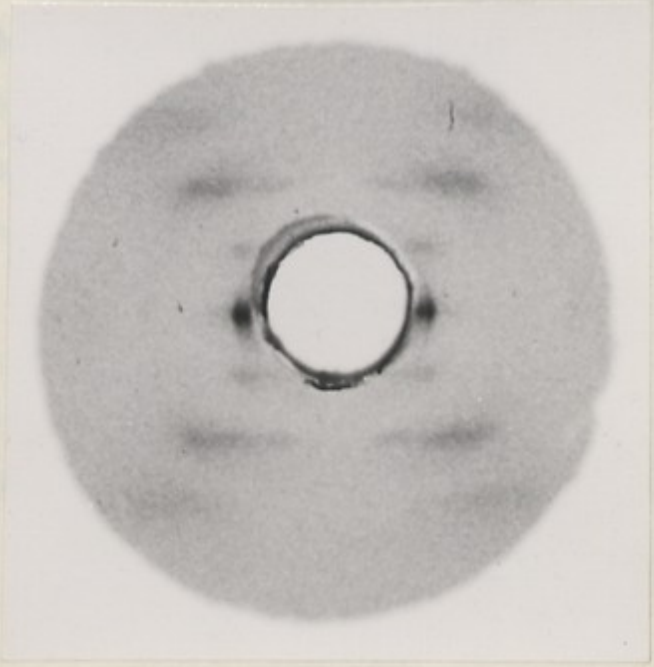


PLATE 10 (b)

Enlarged centre of Plate 6(a), showing equatorial doublet.

fibre but causes no further appreciable length change, or change in the general form of the X-ray diagram. It would seem, therefore, that Structure B may quite possibly be close to the molecular configuration of NaDNA in solution. The titration work of Gulland and Jordan, also Lea and Peacocke, mentioned in Chapter II has suggested that in aqueous solution at neutral pH small stable aggregates of polynucleotide chains exist, linked by hydrogen bonds between their base groups and having their phosphate groups exposed. Since the transformation A to B is readily and rapidly reversible, it seems reasonable to suppose that these small molecular aggregates of the wet state can be easily derived from the grouping existing in Structure A. This suggests that in the crystallites we may expect to find a small group of chains held together as a unit by hydrogen bonds between their base groups and that these units are linked in three dimensional order by phosphate-phosphate bonds of the type postulated above.

(iii) Conclusions.

It may be concluded that, in general, the NaDNA structure consists of hydrogen bonded units of polynucleotide chains linked by phosphate-phosphate bonds and associated water molecules. Furthermore, that these polar phosphate groups lie on lines or surfaces which extend throughout the structure. Hydration, swelling and solution will then result from the distention of these polar regions leaving the hydrogen bonded units intact, but not necessarily always in the same configuration.

This conception of polynucleotide chains forming units in which the hydrogen-bonded groups are turned inwards and the phosphate groups outwards, explains the ready availability of phosphate groups for interaction with proteins. Moreover, since both the protein in nucleo-proteins and the water in NaDNA fibres are believed to be associated with the phosphate groups, it is conceivable that the NaDNA structure might exist without great modification in the nucleo-proteins, the protein taking the place of the structural water. That this would appear to be true has been shown by the similarity between the diffraction patterns of nucleo-protein and Structure B, obtained from various specimens by Wilkins and Randall.[9]

The molecular configuration of the hydrogen bonded units, mentioned above, is discussed in the following Chapters, dealing with the quantitative analysis of the X-ray diagrams.

REFERENCES IN CHAPTER IV.

1. Riley, D.P. and Oster, G. (1951): Biochim. et Biophys. Acta. 7. 526.
2. Astbury, W.T. and Bell, F.O. (1938): Cold Spring Harb. Symp. Quant. Biol. 6. 109.
3. Astbury, W.T. (1947). S.E.B. Symp. 1. 66.
4. Wilkins, M.H.F., Gosling, R.G. and Seeds, W.E. (1951). Nature. 167. 759.
5. Wilkins, M.H.F. (1953): Private Communication.
6. Cochran, W., Crick, F.H.C. and Vand, V. (1952). Acta. Cryst. 5. 501.
7. Stokes, A.R., (1951). Unpublished.
8. Drushel, W.A., and Felty, A.R. (1918). Chem. Zbl. 89. 1016.
9. Wilkins, M.H.F. and Randall, J.T. (1953): Biochim et Biophys. Acta. 10. 192.

CHAPTER V

INTERPRETATION OF THE DIFFRACTION PATTERN OF STRUCTURE B.

Values of R-space Co-ordinates.

For the measurement of the position of reflections the micro-photographs were projected on a white Bristol board screen using a magnification of about 10. The centres of the reflections were then marked on the card and their x and y co-ordinates measured. For broadened reflections an indication of the length of the arc, or layer line streak, was also recorded. The use of the projector rather than a travelling microscope was found not only to be much less fatiguing, but also to provide a more reliable estimate of the positions of weak reflections, and a more convenient method of making measurements on curved layer lines.

Calibration of the scale of the projected photographs was made by indentification of one or more of the reflections measured, with reflections on photographs taken with a Unicam single-crystal camera. In this way the very strong meridional arc of Structure B was found to correspond to a spacing of 3.31A°. Using this value, the appropriate demagnification factor was calculated for converting the measured x and y co-ordinates of the projection to the scale of a Bernal Chart for a flat film, with a 10 cm specimen to film distance. Hence the average ζ value for each layer line was found and from ~~this~~ the mean value of $\zeta_{1/1}$, the layer line period C_1 was established. Once this had been done the ξ values were obtained from the radial distance, r, of the reflections from the origin. For r and z (z = apparent specimen to film distance of the projection) gave $\rho = \lambda/d$ and $\rho^2 = \zeta^2 + \xi^2$.

TABLE I

The Structure B pattern shown in Plate 4 Chapter IV was calibrated by comparison with several Structure B patterns obtained with the Unicam single crystal camera. The polar arc of Plate 4 was shown to correspond to a spacing of 3.31A°. In this way the specimen to film distance of Plate 4 was calculated and the following ξ and ζ values were derived.

ζ Values. (Correct to ± 0.001.)

L	1	2	3	4	5	6	7	8	9
ζ/L	.046	.046	.046	Absent	.046	Weak	.046	.047	.047

Arithmetic mean of ζ/L = 0.046 ± 0.001 (i.e. C = 33.5A°).
To the accuracy of these measurements the mean value of ζ/L may be taken as 0.0465, corresponding to a layer line period of 33.1A°. i.e. C = 33.1A°

ξ Values.

The following ξ values were calculated from the measured radial co-ordinates of the centres of the observed reflections, assuming the layer-line period to be 33.1A°.

ξ	dA°	L = 1;-		L = 3;-		L = 7;-	
L = 0;-		.035	26.4	.072	9.8	.200	4.03
.0295	52	.053	21.8	.098	9.03	L = 8	
.066	23.2	.066	19.0	.136	7.91	.134	3.89
.082	18.8	L = 2;-		L = 5;-		L = 9;-	
.117	8.7	.037	15.4	.169	5.35	.137	3.50
.281	5.5	.071	13.1	.278	4.24	L = 10;-	
.400	3.85	.098	11.4	L = 6;-		0	3.31
				.239	4.19		

The ξ and ζ values obtained in this way from the projection of the photograph shown in Plate 4 Chapter IV are listed in Table 1. The corresponding reciprocal lattice diagram is shown in Fig. 1. Within the error range of the measurements it was found that the layer line period was ten times the spacing corresponding to the meridional arc. Several photographs were taken with the fibre axis inclined to the X-ray beam, but it was found that this 3.31A° arc could not be resolved into discrete reflections, but always appeared on the meridian. Therefore this reflection was assigned zero ξ value and ζ value $10\lambda/c$. The fact that it appears with the fibre (rotation axis) at right angles to the beam must be due to disorientation of the diffracting units. This would have the effect of replacing the R-lattice point by an arc centered at the origin, the tail ends of this are cutting the reflecting sphere. The apparent ζ value of this reflection would then be $\rho \cos \theta$, which is less than the correct ζ value. This was found to be the case; the meridional arc was just below the level of the 10th layer line on the measured projection. The value of ζ observed was $0.455 \pm .002$ compared with $\rho \cos \theta = 0.452$, confirming that if the reflection has $\xi = 0$ then it does occur on the 10th layer line.

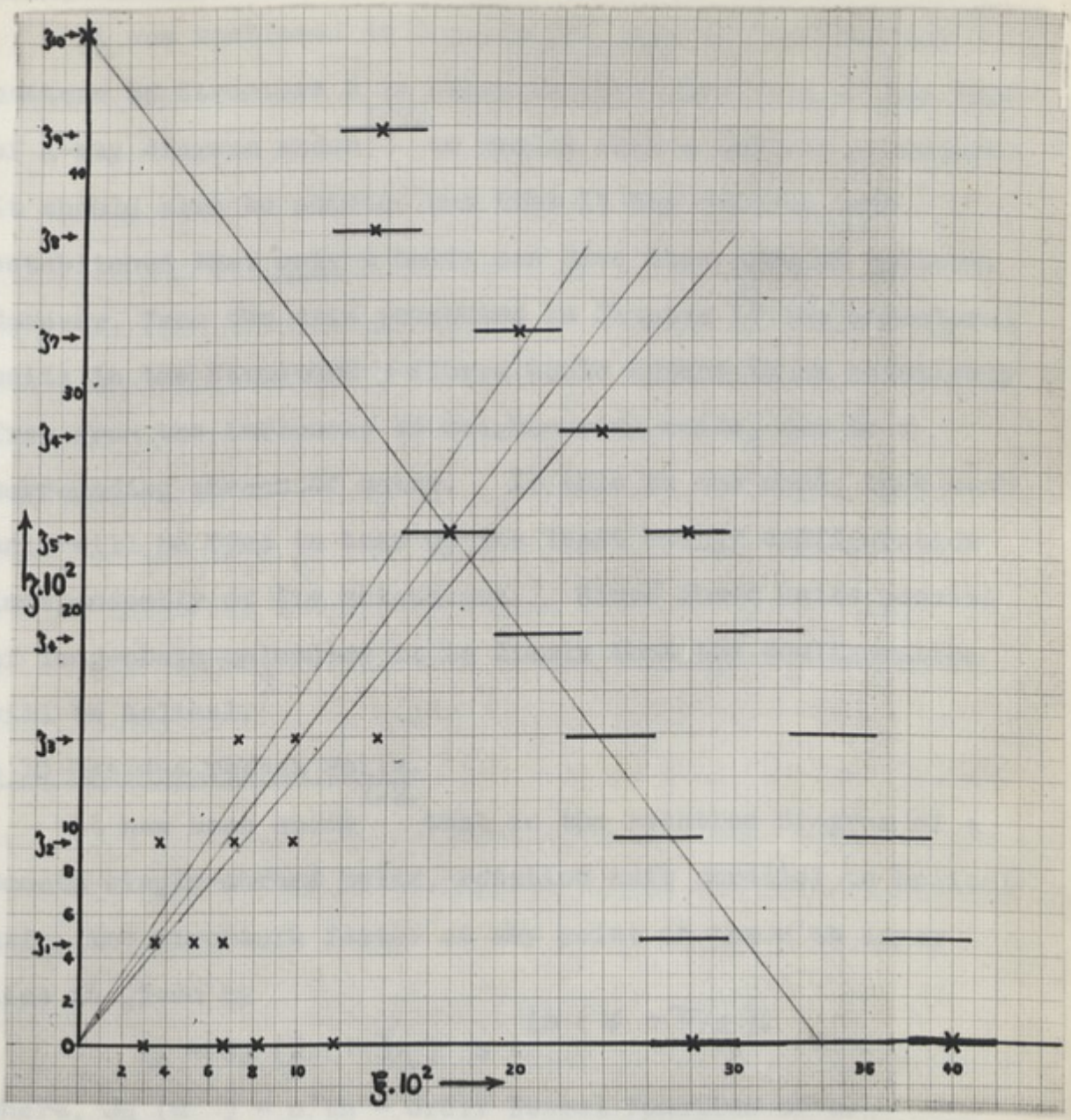


Fig. 1. Reciprocal Lattice diagram of the Structure B pattern shown in Plate 4, Chapter IV.

x = Maximum intensity of reflection.

——— = Diffuse intensity.

Helical Structure.

It was mentioned in Chapter IV, that the diffraction pattern of Structure B is strongly characteristic of the type of X-ray diagram shown[1.2.] to result from a helical structure. It should also be pointed out that it has not yet been established that only a helix can give this type of pattern. However, from the data presented in Chapter IV the structural units in the Structure B fibres would appear to be relatively free from the influence of neighbouring units, due to a surrounding sheath of water. If this is the case, then each unit will be free to take up its least energy configuration independently of its neighbours. Since these units consist of long-chain molecules it is likely that the configuration will be helical.[3]

A 10 Residue repeat Helix.

It has been shown[1.2.] that in the rotation diagram of a smooth single-strand helix, rotation axis parallel to helical axis, the structure factor at any point on the n'th layer line is given by

$$F_n = J_n (2\pi . r . \xi/\lambda) \; e^{in(\varphi + \pi/2)}.$$

where, $J_n(u)$ = n'th - order Bessel Function of u.

r = Radius of the helix.

φ = Azimuthal co-ordinate in R-space.

ξ = Distance from helical axis, in R-space.

λ = Wavelength of incident radiation.

This gives rise to a diffraction pattern in which the innermost maxima on successive layer lines lie approximately on straight lines radiating from the origin, so forming a central

cross. If instead of a continuous smooth helix the structure is formed by a series of residues equally spaced along the helix, then the expression for the structure factor becomes much more complicated. However, if there is a whole number, m, of residues per turn of the helix, the transform is that for a smooth helix with the addition of the central cross pattern repeated with its origin at heights mc^*, $2mc^*$, etc; c being the fibre-axis repeat period. Thus the double diamond pattern, always present in Structure B X-ray diagrams, is characteristic of a helical structure having a whole number of residues per turn of the helix. In this case the second origin is at the meridional arc on the 10th layer line. Lines of maxima radiating from this 3.31A° reflection, as from the origin, may be seen on the fifth and lower layer lines (Fig.1). The two series cross on the fifth layer line, giving rise to the first fairly strong reflection seen at that level. Therefore, if this structure is helical, there are 10 residues per turn of the helix.

Radius of Helix formed by the Phosphate Groups.

Since the linear array of innermost reflections is one of the strongest features of the X-ray diagram, it may be concluded that the outside of the diffracting unit is a helix formed by residues which are, crystallographically, important parts of the structure. Now the heaviest scattering components in NaDNA are the sodium phosphate groups. This interpretation of the X-ray diagram on a helical basis, therefore, provides further evidence that the phosphate groups are on the outside of the structural units. By applying the theory

mentioned above,[1,2] it is possible to calculate the radius of this outermost helix from the ξ values of the first reflections on the successive layer lines, $\ell = 1$, $\ell = 2$, $\ell = 3$, etc., which should correspond to the first maxima in first, second, third and higher order Bessel functions respectively. Consider the first, rather broad, reflection that occurs on the 5th layer line at the point of intersection of the Origin and 10th layer line series of maxima. If the diffracting unit consists effectively of a single helix the ξ value of this reflection is related to the radius of the helix by the expression:-

$6.42 = 2.\pi.r.\xi/\lambda$. where r = Radius of helix

$\lambda = 1.539A^{o}$

6.42 = Value of u at the first maxima of $J_5(u)$.

Now the ξ value of the centre of this reflection is 0.169 and hence the corresponding value of r is $9.3A^{o}$. In Fig.1, the line joining the centre of this reflection to the origin is seen to pass through the first group of reflections on the lower layer lines. However, as may be seen from Plate 4, Chapter IV, the most intense reflections in these lower layer line groups are the 2nd and 3rd reflections in each case. Therefore, it would appear likely that the phosphorus atoms lie on a helix of radius less than $9.3A^{o}$. Unfortunately the detail in the diagram is not sufficient to permit an exact evaluation of this radius. This may be due, in part, to the spatial spread of the phosphate residue. The limits of the range of possible values of r are given by the edges of

the first fifth layer line reflection. These are marked in figure 1 and correspond to values of r equal to 10.7 and 8.3A°. The intensity distribution, mentioned above, of the first three reflections on the 1st, 2nd and 3rd layer lines, favours a value toward the lower limit. It would appear then, that the heaviest scattering members of the phosphate residues, i.e. the Phosphorus atoms, must lie on a helix whose radius is something between 9 and 8.3A°.

The Number of Polynucleotide Chains in the Helix.

Structure B does not occur alone, in fibres showing the A to B change, at humidities lower than 92% R.H. It is likely then, that whatever the humidity, the amount of water to be associated with polynucleotide chains having a well defined Structure B form is greater than that associated with Structure A, and is probably not very far from that found for pellets of NaDNA at 92% R.H. i.e. 60% of the dry weight of the specimen. This means a value of 528 for the molecular weight of a nucleotide with associated water. The density (ρ) of the diffracting material is more difficult to ascertain, since Structure B can co-exist with a large quantity of water. It is likely, however, that at humidities lower than 92% the B helices are close packed and the density quite high, even close to the value 1.471 g/cc of the crystalline A form, since the two states can co-exist as shown by mixture diagrams. Therefore, for the purpose of approximate calculation, we may assume a hexagonal closely packed structure of density 1.47 g/cc, with the diffracting unit as a cylinder of radius 8.5A°. The volume of the hexagonal primitive cell (a = 17.0A°, c = 33.1A°)

is then the minimum volume to be associated with one B helical unit of NaDNA. Hence an approximate value for x, the number of nucleotides (i.e. phosphate residues) in one turn of the helix may be derived:-

$$x = \frac{1.47.17^2.33.1.0.8660}{528 \quad 1.66} = 20.4$$

NB if $\rho = 1.38$ g/cc, $x = 19.1$

Since the X-ray diagram requires the helical units of Structure B to have a repeat of 10 phosphate residues per turn of helix, the above calculation suggests that each unit contains phosphorus atoms on <u>two</u> co-axial helical chains with a common radius of about 8.5A°. If these two chains were spaced equally along the fibre axis, the diffraction theory would require the n'th layer line to depend on $J_{2n}(u)$, where u is $2\pi.r.\zeta/\lambda$ as defined above. This would mean that the first reflections observed on the successive layer lines 1, 2, 3 etc., should correspond to the first maxima of the functions $J_2(u)$, $J_4(u)$ $J_6(u)$ etc., respectively. The linear array of the origin series of reflections would then indicate a value of about 17A° for the radius of the helix, which leads to an impossibly low value for the density. Therefore the chains must be unequally spaced along the fibre axis. If one strand is displaced relative to the other by about three-eighth's of the fibre axis period, i.e. 13.2A°, it would account for the absence of the fourth layer line and the weakness of the sixth in the X-ray diagram.

<u>Packing Of Helices.</u>

The equatorial reflections in Plate 4, Chapter IV, are not those expected from a hexagonal lattice. Furthermore, in no

Structure B pattern, at any humidity, has the ratio of equatorial reflections been found to correspond to that of hexagonally packed units. However, it is found that the fairly sharp reflections observed on the equator and first three layer lines may be indexed in reasonable agreement, as shown in Table 2, with a monoclinic cell having the following parameters.

$a = 26.2A^{o}$, $b = 52A^{o}$, $c = 33.1A^{o}$, and $\beta = 97\frac{1}{2}^{o}$.

Now Plate 4 was obtained at 75% R.H., therefore the density of the crystallite regions may be close to that found for pellets of NaDNA at 75% i.e. $\rho = 1.471$ g/cc. Fibres at this humidity show about 40%-45% water up-take (percentage of dry weight) but the water content of a Structure B crystallite is likely to be higher than that, possibly 60%, as was suggested above. Thus if k is the number of two-strand helices running through each unit cell then:-

$$k = \frac{1.471.\ 26.\ \ 52.\ \ 33.1}{20.528.\ 1.66} = 3.8$$

Here the molecular weight of a "wet" nucleotide, M, has been taken as 528, i.e. 60% by weight of water, corresponding to 11 molecules of water per nucleotide. If M were equal to 474 i.e. 43.6% water, then the value of k would be 4.2.

Non-Circular Cross-Section.

The argument presented above indicates that the proposed unit cell probably has 4 two-strand helical molecules passing through it. One simple way of packing them is shown in Fig.2. This flattened hexagonal arrangement suggests that the helices are not of circular cross-section. The fact that a complete halving of the unit cell does not occur could then be explained

TABLE II

In this table observed ρ/ρ_0 values are compared with those calculated on the basis of a monoclinic cell having $a = 26.2A^o$, $b = 52.0A^o$, $c = 33.1A^o$ and $\beta = 97.5^o$.

hkl	ρ/ρ_0 calc.	ρ/ρ_0 obs.
010	.0295	.0295
100) 020)	.059	Weak or Absent
110	.066	.066
120	.083	.082
300) 060)	.177	.177
011	.031	.035
101	.0529	.053
101 111	.065 .072	.066

hkl	ρ/ρ_0 calc.	ρ/ρ_0 obs.
012	.032	.037
102	.071	.071
122 $\bar{1}32$	.083 .101	.098
103 $\bar{1}23$	.076 .072	.072
123 $\bar{1}33$ 203	.097 .099 .0995	.098
203 $\bar{2}33$ 213	.136 .135 .139	.136

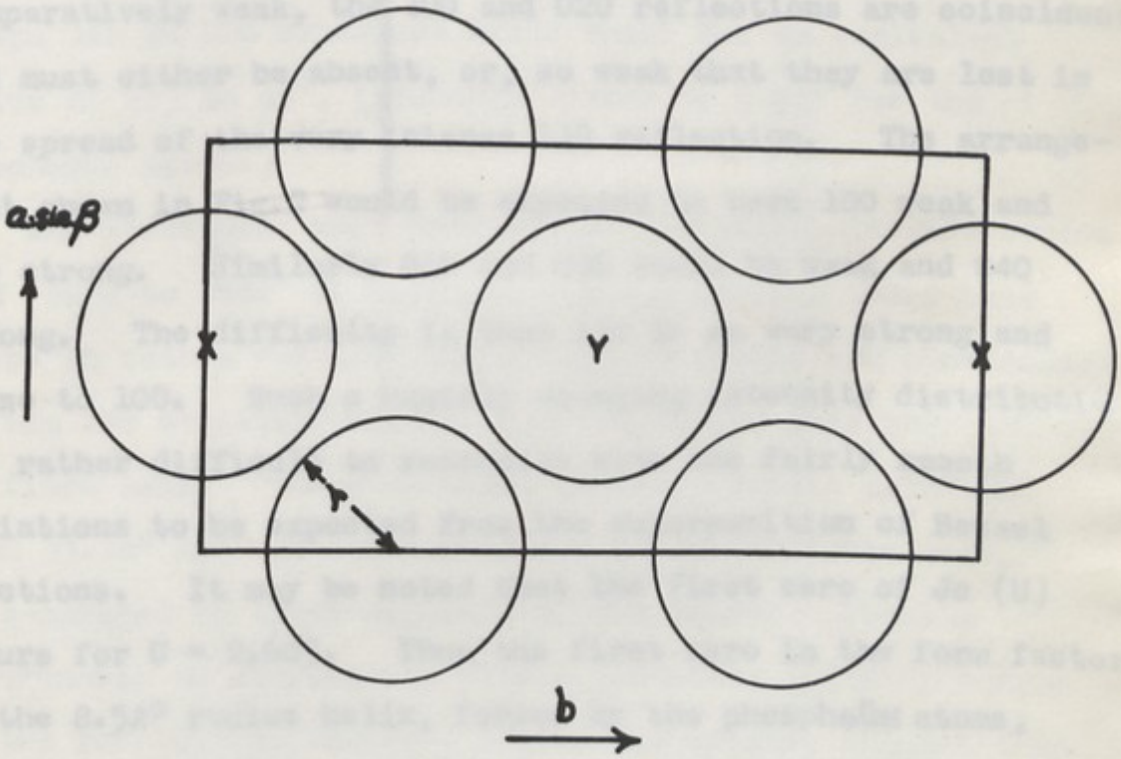


FIG.2: PACKING OF STRUCTURE B HELICES.

UNIT CELL :-

$a = 26{\cdot}2\,A^{\circ}$; $b = 52{\cdot}0\,A^{\circ}$; $r = 8{\cdot}5\,A^{\circ}$. ($c = 33{\cdot}1\,A^{\circ}$)

$\beta = 97{\cdot}5^{\circ}$.

SCALE. 2mm. = $1\,A^{\circ}$.

by molecules X and Y (Fig. 2) having different orientations of their cross-sections.

Equatorial Intensity Distribution.

Although the suggested unit cell accounts quite well for the 1st, 2nd and 3rd layer line reflections, the equatorial indexing raises some difficulties. The 010 reflection is comparatively weak, the 100 and 020 reflections are coincident and must either be absent, or, so weak that they are lost in the spread of the very intense 110 reflection. The arrangement shown in Fig.2 would be expected to have 100 weak and 200 strong. Similarly 010 and 020 would be weak and 040 strong. The difficulty is that 110 is so very strong and close to 100. Such a rapidly changing intensity distribution is rather difficult to reconcile with the fairly smooth variations to be expected from the superposition of Bessel functions. It may be noted that the first zero of Jo (U) occurs for U = 2.605. Thus the first zero in the form factor of the 8.5A° radius helix, formed by the phosphorus atoms, occurs at $\xi = 0.075$. If the rest of the molecule is considered as a series of ~~concentric~~ co-axial helices the zero's in their respective form factors will all fall at ξ values larger than that due to the phosphorus helix. Therefore the 100 and 020 reflections, if dependent only on the form fact due to the helical molecule, should be more intense than the 110 reflection. Hence, if the indexing is correct, the large difference in the observed intensity values of 100 (020) and 110 must be due to the departure of the helices from circular cross-section, (probably an elongation near to the ~~100~~ 110 direction).

Conclusions.

Whilst a complete interpretation of the fibre-diagram of Structure B has not been attempted, the following general conclusions emerge from the calculation and discussion presented above.

The molecule of DNA is probably helical. The phosphate groups lie on the outermost helix which has an equivalent radius of 8.5 to 9A°, (probably nearer to 8.5A° for the phosphorus atoms.)

The structural unit consists of two polynucleotide chains held "back to back" by hydrogen bonding between their base groups in such a way that the phosphate groups are constrained to form two co-axial helices with the common axis parallel to the fibre axis. These co-axial helices, radius 8.5A° and pitch 33.1A° are not equally spaced along the fibre axis, but separated by three-eighth's of the pitch. Density measurements in conjunction with a tentative indexing scheme for the reflections observed around the origin, suggest that the helices are not of circular cross-section although the departure is probably not large.

REFERENCES IN CHAPTER V

1. Cochran, W., Crick, F.H.C., and Vand, V. (1952): Acta. Cryst. 5. 501.
2. Stokes, A.R. (1952): Unpublished.
3. Pauling, L., Corey, R.B., and Bransom, H.R. (1951): Proc. U.S. Nat. Acad. Sci., 37. 205.

CHAPTER VI

THE CALCULATION AND INTERPRETATION OF THE CYLINDRICAL PATTERSON FUNCTION OF STRUCTURE A.

The fibre diagram reproduced in Plate 1 Chapter IV shows 66 independent reflections distributed on nine well-defined layer lines. As mentioned in Chapter III, attempts to index this pattern by direct inspection had led to rather uncertain results due to the ambiguities of indexing the reflections at large angles, although it was fairly clear that the unit cell was most probably C-face centered monoclinic with the C-axis in the fibre axis direction. Because of this uncertainty in the indexing it was decided to calculate the cylindrical Patterson Function as outlined by MacGillavry and Bruins in 1948.[1] This function does not require the indices of the reflections other than their layer-line numbers. It is, therefore, periodic in the C direction only. It is the function that would be obtained by rotating the full three-dimensional Patterson around the fibre axis, and then taking the average axial section. The a and b lattice translations should be revealed as important peaks in this Cylindrical Patterson. Therefore, it was felt that calculation of this function would not only supply structural information, but also resolve the uncertainties in the indexing and enable the full three-dimensional, with higher information content, to be calculated.

Measurement of R-Space Co-ordinate z.

The projection method was used, as described in Chapter V. The calibration reflection in this case was the very strong second equatorial reflection. This was shown, by photographs taken with the Unicam single-crystal camera, to

correspond to a spacing of 11.3A°.

Correction to ζ for Tilting of the Fibre Axis.

It was almost impossible to place the fibre on the collimator face so that it was accurately perpendicular to the X-ray beam. In fig.1. (a) the diffraction conditions are illustrated for a fibre inclined at $(90 - \gamma)$ to the direct beam. A plane, in R-space, parallel to the layer planes and at a distance from the R-space origin O given by $\zeta_s = \sin\gamma$ would give a straight line l_s on the flat film. Therefore layer planes above ζ_s, would form layer lines which were less curved and had reflections more widely separated than if the fibre were normal to the direct beam. Layer planes lying below ζ_s would form layer lines more highly curved having reflections closer to the meridian. The general form of the layer lines would be as indicated in Fig.1. (b).

Let ζ_1, and ζ_2 be the apparent ζ- values measured above and below the equator respectively, for a given set of reflections corresponding to a spacing $d = \lambda/\rho$. It can be shown that the required value of for the perpendicular fibre is given by

$$\zeta = \zeta_1 \cdot \cos\gamma + \tfrac{1}{2}\rho^2 \sin\gamma = \zeta_2 \cos\gamma - \tfrac{1}{2}\rho^2 \sin\gamma \quad (1)$$

whence $$\tan\gamma = \zeta_2 - \zeta_1 / \rho^2 \cdot \quad (2)$$

and $$\zeta = \cos\gamma \cdot (\zeta_2 + \zeta_1)/2 \cdot \quad (3)$$

Measurements of ζ_1, ζ_2 and ρ can, therefore, be used to determine γ .

Most of the measurements were made on photographs for

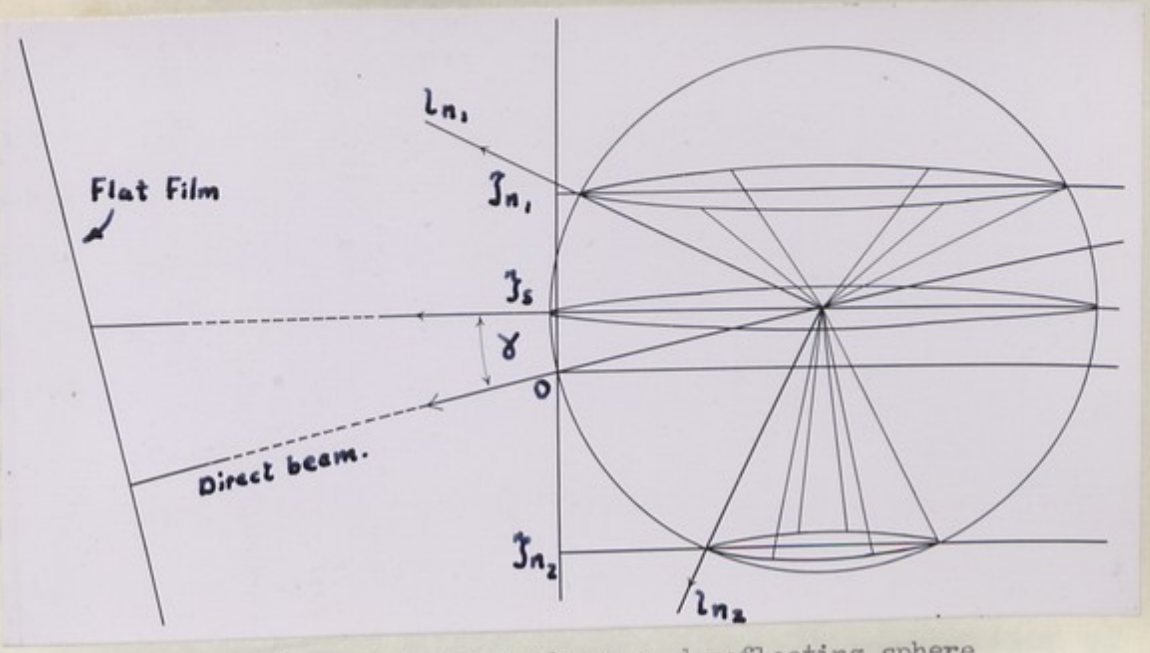


FIG.1 (a) Reciprocal lattice planes and reflecting sphere for a tilted fibre

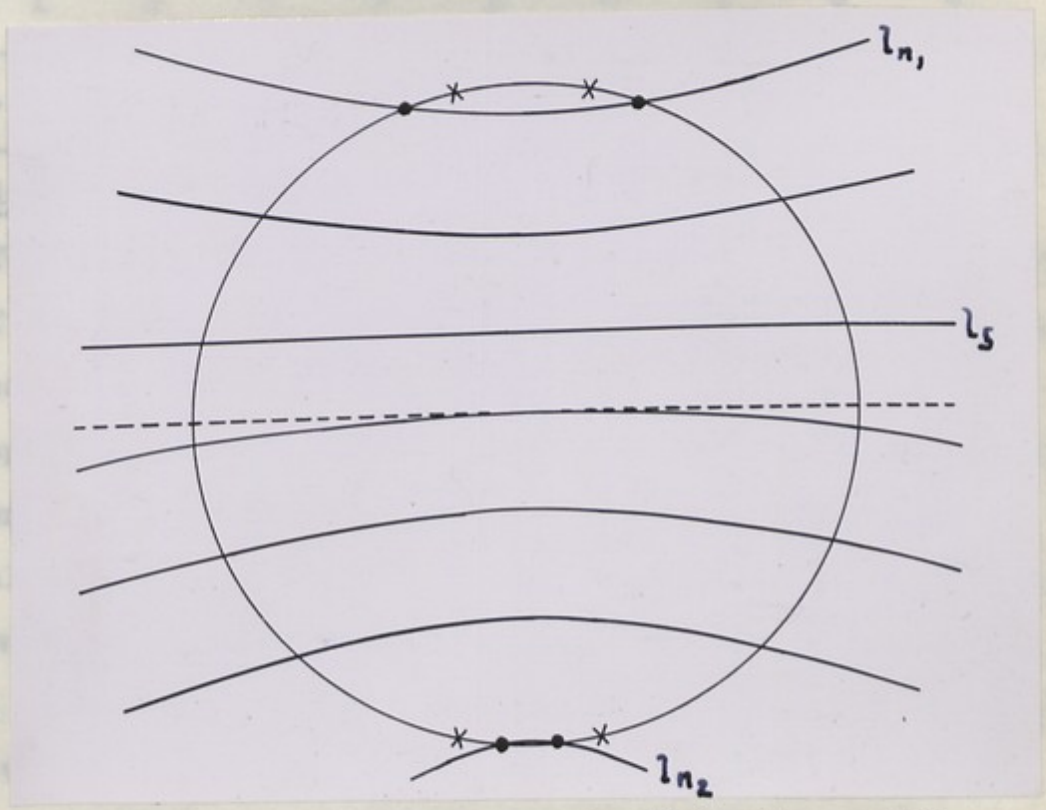


FIG.1 (b) Appearance of photograph for diffraction conditions shown in (a)

X ≡ Position of equivalent reflections with fibre normal to the X-ray beam

• ≡ Position of equivalent reflections for fibre inclined at (90 - γ) to the X-ray beam

which γ was found to be about 8°. Values of ζ, obtained with the aid of equations (1) and (2) from measurements of reflections on all the nine layer lines, are shown in Table 1. Each value is a mean of measurements on at least 3 reflections. Neglecting the value obtained for the 3rd layer line, where the reflections were weak and difficult to measure accurately, the mean value of ζ is 0.0547. This gives c, the fibre-axis repeat period, as 28.1A.

TABLE I

L	1	2	3	4	5	6	7	8	9
ζ/L	.0550	.0540	.0533	.0545	.0552	.0548	.0545	.0546	.0551

Calculation of ξ Values.

The most accurately observable parameter of any reflection in a fibre diagram is the radius r of the circle on which the four equivalent photographic spots lie. Therefore, using the apparent specimen film distance calculated from the 11.3A° "calibration" reflection, the value of ρ $(= \lambda/d)$ was calculated for each reflection from it's observed value of r. Once the layer line spacing, ζ, had been established ξ values were then determined from the relationship $-\xi^2 = \rho^2 - \zeta^2$. The ξ values calculated in this way are listed in Table II.

Measurement of Intensities.

The variety of shapes and sizes of the photographic spots made it impossible to estimate directly the integrated photographic intensity. It was necessary, therefore, either to explore each spot photometrically, or to estimate its maximum

intensity and to consider separately the question of spot size and shape. The latter alternative was adopted, since in the micro-photographs available the grain-size was too large in relation to the spot-size for photometry to be applied with any degree of accuracy. The use of fine-grain films would have lead to excessive exposure times, when using single fibres, and thus removed the advantage gained from the use of the micro-technique.

The maximum intensity of each spot was estimated visually by comparison with a standard scale. For the preparation of this scale the ultra-fine collimating system of a low-angle camera* (slit width 8μ) was used to photograph the direct beam, with exposure times varying from 3 to 90 seconds. A set of streaks of width comparable with that of the spots on the micro-photographs was obtained. The fibre-photograph was projected, as for the measurement of the positions of reflections, and the scale displaced by hand across the photograph in the projector. The standard streaks were placed as close as possible to the spot whose intensity was being measured in order to eliminate as far as possible, the effect of varying intensity of the background.

Relation between Maximum Intensity and Integrated Intensity.

The relation between the measured photographic maximum intensities of the reflections and their R-space integrated intensities depends on the following factors:-

(i) True diffraction breadth

The breadth of the reflections, measured along a line joining the reflection to the origin, is approximately constant

* Designed by K.P. Norris of this laboratory.

over the whole photograph. It may, therefore, be assumed that the true diffraction breadth is everywhere small compared with the geometrical broadening (direct beam size).

(ii) Spread of reflections in R-space due to disorientation of the crystallites.

(a) If the crystallites are perfectly aligned with respect to the fibre axis but in random orientation about that axis, each reflection will be spread over a circle of radius ξ/λ. Maximum intensities must, therefore, be multiplied by $2\pi\,\xi/\lambda$ in order to relate them to integrated intensities.

(b) Imperfect alignment of the crystallites with respect to the fibre axis results in a small angular spread, ϕ, (fig.2) of each reflection along an arc of a circle with centre at the origin and the fibre axis as diameter. The length of this arc in R-space is equal to $s\,\phi$, where s is the radial distance of the arc from the origin, ($s = 2(\sin\theta)/\lambda$). Since ϕ is a constant the R-space maximum intensity need only be multiplied by s to relate it to the integrated intensity.

(iii) Geometrical factors involved in the transfer of R-space effects to the photographic film.

Each reflection will be spread, in R-space, over a small volume of circular symmetry with mean radius ξ/λ, and axis on the fibre axis. An axial section of this volume (Fig.2) shows two elements, each of area ds.dl; ds is the true diffraction breadth of the reflection and dl. ($= s\,\phi$) is the length of the arc resulting from disorientation of the crystallites with respect to the fibre axis. The effect on the photographic

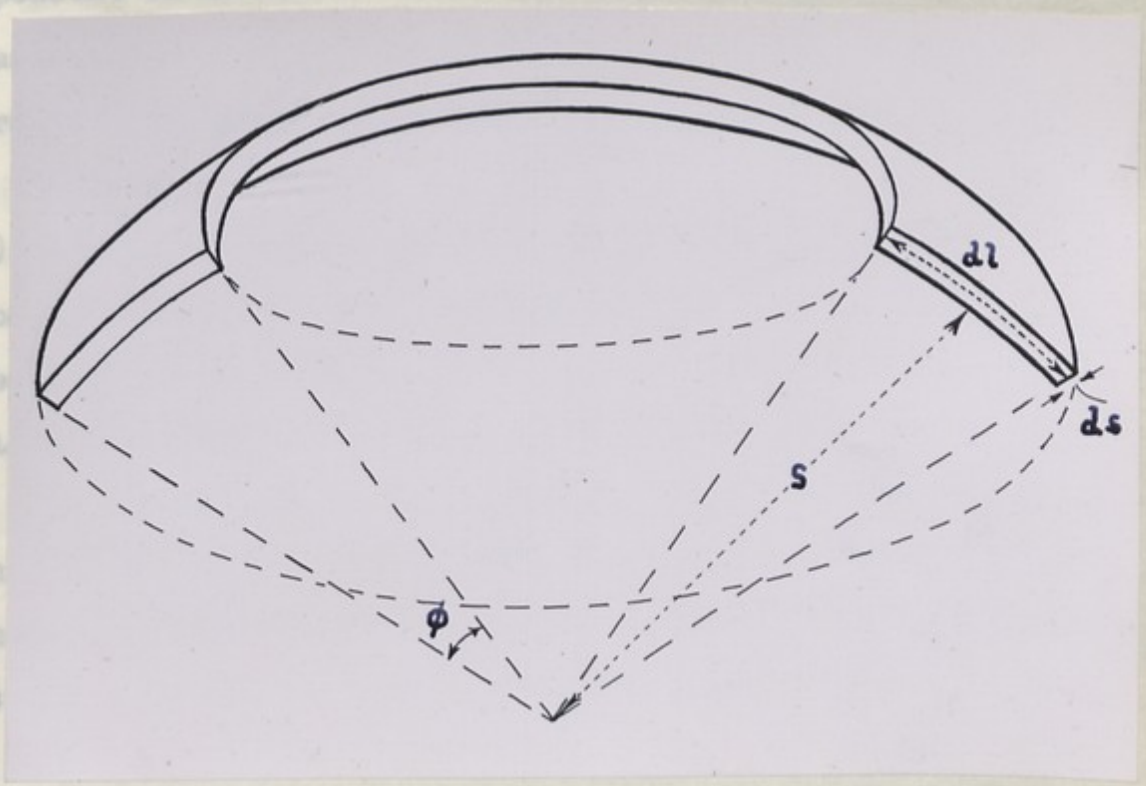


FIG. 2 Diagram illustrating spread of reflection in rotated reciprocal space .

maximum intensity of the oblique intersection of ds and dl with the reflecting sphere must, therefore, be considered.

When a reflection intersects the reflecting sphere obliquely two extreme effects are possible.

(a) If the diffraction breadth is much less than the geometrical breadth then the photographic spot-size is independent of the diffraction breadth, and oblique intersection leads only to increased photographic intensity. This is normally true in single-crystal work and is corrected for in the Lorentz factor.

(b) If the diffraction breadth is much greater than the geometrical breadth then oblique intersection with the reflecting sphere leads to increased spot breadth on the film but does not influence the maximum intensity.

In the present case (a) is true for the dimension ds. Since ds makes an angle θ with the surface of the reflecting sphere, the observed maximum intensity must be multiplied by cos. θ to relate it to the integrated intensity.

With regard to the dimension dl, it happens that all reflections for which the intersection of dl and the reflecting sphere have appreciable obliquity occur at diffraction angles large enough for the length of the photographic arc to be much greater than the geometrical broadening (direct beam size). Therefore condition (b) holds true and no special correction is required.

At small diffraction angles, where the diffraction broadening (dl) is not very much larger than the direct beam

size, some correction is necessary for the effect of the geometrical broadening on the apparent photographic maximum intensity. Now dl is given by $S\phi$ where ϕ is the angular spread due to disorientation. To obtain ϕ a visual estimate was made of the half-peak length of the arcs for equatorial reflections at large θ. Since the value of the correction to be applied was in any case small, this method of estimation was considered adequate. This value of ϕ was then used to calculate the R-space lengths, dl, of arcs occurring at small θ. The lengths, dl', of the photographic arcs for reflections at small θ were then obtained, by calculation, from the dl values and the known diameter of the direct beam using the method of Jones.[3]

If the effect of oblique intersection with the reflecting sphere be neglected, the ratio of the R-space maximum intensity to the observed photographic maximum intensity is then given by dl'/dl. The measured maximum intensity values of reflections at small angles must therefore be multiplied by this factor.

It is shown in what follows that any error introduced by neglecting the obliquity correction for the reflections at small θ has no serious effect on the resulting Patterson function.

Intensity Corrections

From the above discussion it follows that the R-space integrated intensity may be derived from the observed maximum intensity values by applying the following corrections:-

1. Multiply by S/λ . (see (ii) (a) above).
2. Multiply by cos θ. (see (iii) (a) above).

3. Multiply by s (see (ii) (b) above).

4. For reflections at small θ, multiply by di/dl.

5. Multiply by the polarisation factor $2/(1 + \cos^2 2\theta)$.

6. Correction for the influence of the angle of incidence of the diffracted X-rays on the film.

7. Correction for variation of specimen-film distance with θ for a flat film.

8. Absorption correction.

Correction 8, was negligible throughout. Moreover, in the range of θ used in this study, corrections 2, 5, 6 and 7 taken together never exceed 10%, which is within the limit of experimental error of the intensity measurements. Only corrections 1, 3 and 4 were therefore used.

The equatorial reflections of any given value of s form only <u>two</u> equivalent photographic spots, whilst the reflections for any value of s with $L > 0$ contribute to <u>four</u> equivalent spots. Therefore the observed intensities of the equatorial reflections must be halved.

For the complete set of intensity measurements, corrected values shown in Table II, twelve films were used. These were obtained from two exposures on each of two different specimens, using three films for each exposure. The total range of intensities measured in this way was from 1 to 96, and agreement obtained between different sets of measurements was better than 20% for all expect a few of the weakest reflections. Some further small error may have been introduced by the approximations involved in obtaining integrated from maximum intensities,

TABLE II

DATA USED IN CALCULATION OF THE CYLINDRICAL PATTERSON SHOWN IN FIGURE 3.

$s/\lambda \cdot 10^2$	I_c	$\frac{1}{2} \cdot I_c \cdot e^{-a \cdot s^2}$
ZERO LAYER		
5.25	10	4.5
8.85	157	67
10.00	4	1.5
10.50	5	2
13.25	19	6.5
13.80	17	5.5
15.75	9	2.5
17.55	17	4.5
18.20	42	10.5
18.85	25	6
20.00	17	3.5
20.80	12	2.5
22.10	33	6
22.85	63	10.5
24.00	46	7
1st LAYER		$I_c \cdot e^{-a \cdot s^2}$
4.90	24	22
5.50	12	11
8.60	46	39
9.05	28	23
10.10	13	10
13.10	24	16
15.30	11	7
16.10	12	7
22.60	28	10
23.50	30	9
2nd LAYER		
4.55	50	42
4.90	50	42
5.90	35	29
9.80	27	20
11.15	27	19
13.00	67	42
17.20	64	31
18.05	59	27
20.00	11	5
21.90	36	12
22.90	51	16
26.60	19	4
27.90	21	3

$s/\lambda \cdot 10^2$	I_c	$I_c \cdot e^{-a \cdot s^2}$
3rd LAYER		
12.80	13	7
14.55	13	7
14.95	16	8
19.80	6	2
21.50	7	2
4th LAYER		
12.35	7	3
14.10	43	19
14.95	64	27
16.90	50	18
21.45	22	5
22.75	24	5
5th LAYER		
12.40	33	25
14.00	9	3
6th LAYER		
8.40	35	12
10.40	92	28
11.90	69	19
13.50	79	21
7th LAYER		
7.90	39	9
11.00	56	11
12.85	59	12
8th LAYER		
10.80	129	16
12.50	133	15
15.70	39	3
9th LAYER		
5.70	105	2

NB Ic = Corrected intensity.
(ie. Integrated intensity)
a = 4.56
s = $2 \sin \theta / \lambda$.

but such error will vary only smoothly and slowly with "s" and will not, therefore, have a major influence on the main features of the Patterson diagram.

Artificial temperature factor.

Since the corrected intensities showed little, if any, tendency to decrease with increasing θ and "artificial temperature factor" was applied. Intensities were multiplied by $e^{-a^2 s^2}$ where the value of a was chosen to be 4.56. This reduced to 0.3 of its value the intensity of the furthest equatorial reflection observed.

Calculation of the Cylindrical Patterson Function.

The cylindrically symmetrical Patterson function $\phi(z, x)$ is given by the equations:-

$$\phi_l(x) = \frac{2\pi}{NV} \int_0^\infty H(l, \xi') . J_0(2\pi . \xi' . x) . \xi' . d\xi' \quad (1)$$

and

$$\phi(z,x) = \sum_l \phi_l(x) . \cos 2\pi l . x . \quad (2)$$

where $\xi' = \xi/\lambda$, ξ being the Bernal Co-ordinate and λ the wavelength of the incident radiation.

$H(l, \xi')$ = The observed, corrected, intensity on the lth layer-line relating to a distance ξ' from the fibre axis in R-space.

$J_0(u)$ = The zero order Bessel function of U.

V = Volume irradiated.

N = Total number of periods in the fibre axis direction.

If each layer-line shows only discrete reflections whose integrated intensities can be measured, the integral in equation (1) may be replaced by a summation thus:-

$$\phi'_l(x) = \sum_{\xi'} I(l, \xi') . J_0(2\pi . \xi' . x) . \quad (3)$$

where I (l, ξ') is the integrated intensity of a reflection on the lth layer-plane in R-space, distant ξ' from the fibre axis. Hence $\phi_l'(x) = N.V.\phi_l(x)$.

The values of the 63 corrected (integrated) intensities, I (l, ξ'), were therefore assembled in groups corresponding to their layer line value l . The product, I (l, ξ').Jo($2\pi\xi'x$), was then calculated at intervals of $1A^{\circ}$ in x over the range $x = 0$ to $50A^{\circ}$. In this way the summation of equation (3) was carried out, and values of $\phi_l'(x)$ per layer line, for each value of x, were obtained. In order to evaluate $J_0(2\pi\xi'x)$ over the desired range in x, values of $J_0(u)$ for u up to 76 were required. Values of $J_0(u)$ were only available for u = o to 40. 4.5. Therefore the higher orders of $J_0(u)$, u = 40 to 76, were calculated from the approximation formula:-

$J_0(u) = \frac{\sin u + \cos u}{(\pi u)^{1/2}}$, at intervals of 0.3 in u. The results obtained are tabulated in the Appendix. For the calculation of $J_0(2\pi.\xi'.x)$ it was found convenient to express the values of $J_0(u)$ above u = 15, graphically. To the accuracy of the intensity measurements a plot of $J_0(u)$ correct to 0.001 was sufficient.

Having calculated $\phi_l'(x)$ the required Patterson function was then obtained, for each value of x, as a function of z from the cosine series of equation (2). The summation was carried out with Beevers-Lipson strips.

The resulting cylindrically symmetrical Patterson function is shown in Fig. 3. The contours were drawn at relative field figure intervals of 20. The dotted lines

indicate negative values.

Influence of Errors in Intensity Measurements.

The greatest possible source of error lay in the two strongest spots, these were the 11.3A° reflection on the equator and the 11.8A° reflection on the 2nd layer line. Each of these was an unresolved doublet, a fact which was allowed for only by making a visual estimate of the breadth of the arc. Moreover, in the case of the 2nd layer line doublet, no correction was made for oblique intersection with the reflecting sphere (see above) and the intensity value adopted was therefore somewhat too high. It was estimated that, for these reflections, the maximum possible error from all sources was about 30%. The cylindrical Patterson function was therefore calculated, for values of x up to 26A°, for these two doublets alone and one-third of it subtracted from the function shown in Fig.3; the result is shown in Fig. 4. The principal features of the function were substantially unaltered by this procedure, so that the effect of the estimated experimental error in the intensity measurements was not serious.

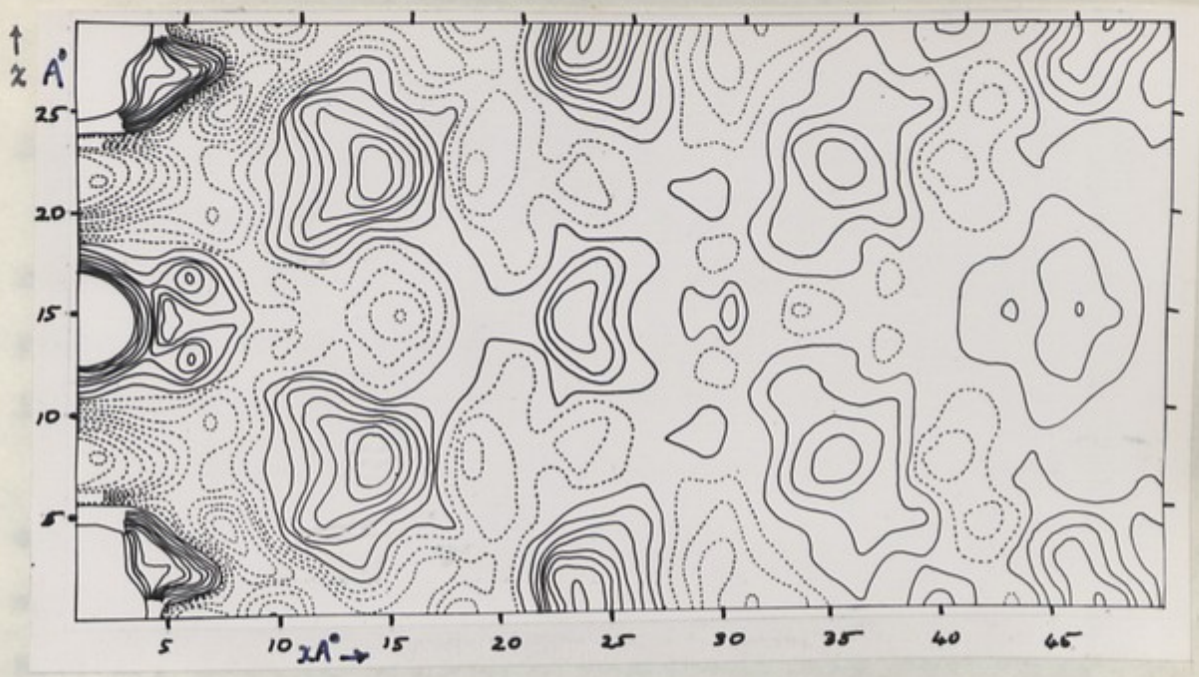


FIG. 3 The cylindrical Patterson function of Structure A.

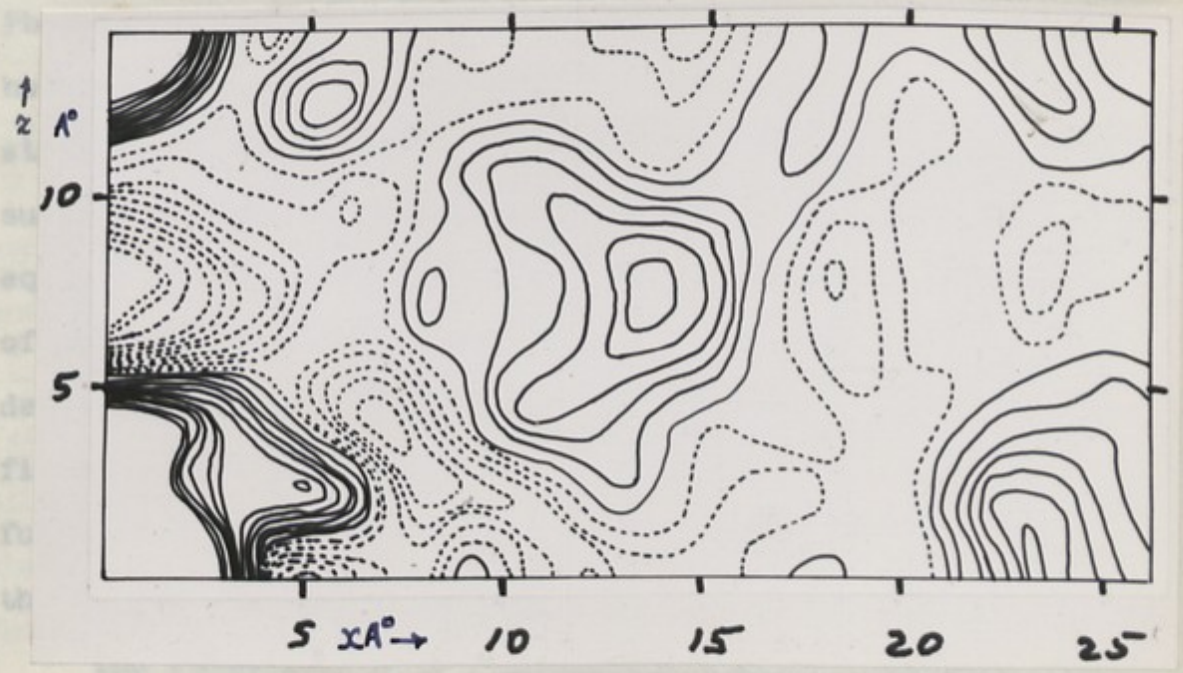


FIG. 4 A half-cell height plot of the function shown in Fig. 3, with the contribution of strong reflections reduced by one-third.

Interpretation of the Cylindrical Patterson.

(i) A Helix in Structure A.

If a co-axial two-chain helical molecule exists in Structure B it ~~must also~~ probably exists although in modified form, in Structure A, since the change from A to B is easily and rapidly reversible. In this case structural deductions might be expected to be more exact because of the far greater amount of information presented by the A pattern.

The Patterson showed rather few, obviously composite peaks of fairly high intensity relative to that of the origin. This was probably due to the small number of reflections available. However, since the unit cell contained a very large number of atoms, circa 1000 (excluding hydrogen and water), it seemed likely that the main features of the Patterson represented Phosphate-Phosphate vectors, the phosphate groups being the heaviest groups in the structure. On this basis the very strong peak on the fibre axis at the half-cell height at once suggested that the two co-axial phosphate chains expected were equally spaced along the fibre axis, with the phosphate groups of one chain exactly above those of the other. In order to determine if these two chains were in helical form, it was first necessary to decide on the general form of the Patterson function of a helix and then to identify such a function in the Patterson of Structure A.

The Patterson of a single, very narrow, helix of continuous and uniform electron density will be a smooth curve of positive vector density. For a helix whose axis is coincident with the fibre axis this curve may be constructed from the

equations:-

$$x = 2.\ r.\ \sin \alpha/2 \qquad \text{and } z = k\alpha.$$

where r = Radius of helix

α = Angular orientation, (projected on a plane at right angles to the fibre axis), of the intra-helical vector having co-ordinates x and z. (in the sense shown in Fig.3).

k = A constant determined by the pitch of the helix

For Structure A the value of z for one turn of the helix is 28.1A°, the height of the unit cell, hence $k = 28.1/2\pi = 4.47_2$.

The presence of a single, narrow helical structure with one turn in the c period, should therefore be shown in the Patterson as a chain of vector peaks lying on a curve whose maximum x value occurs at the half-cell height and is equal to the diameter of the helix. Between this chain of peaks and the c axis there should be low vector density and a marked absence of vectors in the c direction or at small angles to c. In Fig.3 a narrow curve of positive vector density occurs around the $z = \frac{1}{2}C$ level with x about 18A°. This suggested the presence of a helix whose radius was 9A° to within quite close limits (.2A°). This is shown more clearly in the corrected Patterson of Fig. 4. Therefore the Patterson of a single continuous helix of radius 9A° was calculated, according to the above equations, and superimposed on the corrected Patterson. This curve is denoted by (i) in Fig.5. To this was added curve (ii) which completed the intra-helical vectors to be expected from two continuous co-axial helices separated by $z = \frac{1}{2}c$.

Curves (i) and (ii) together follow most closely the

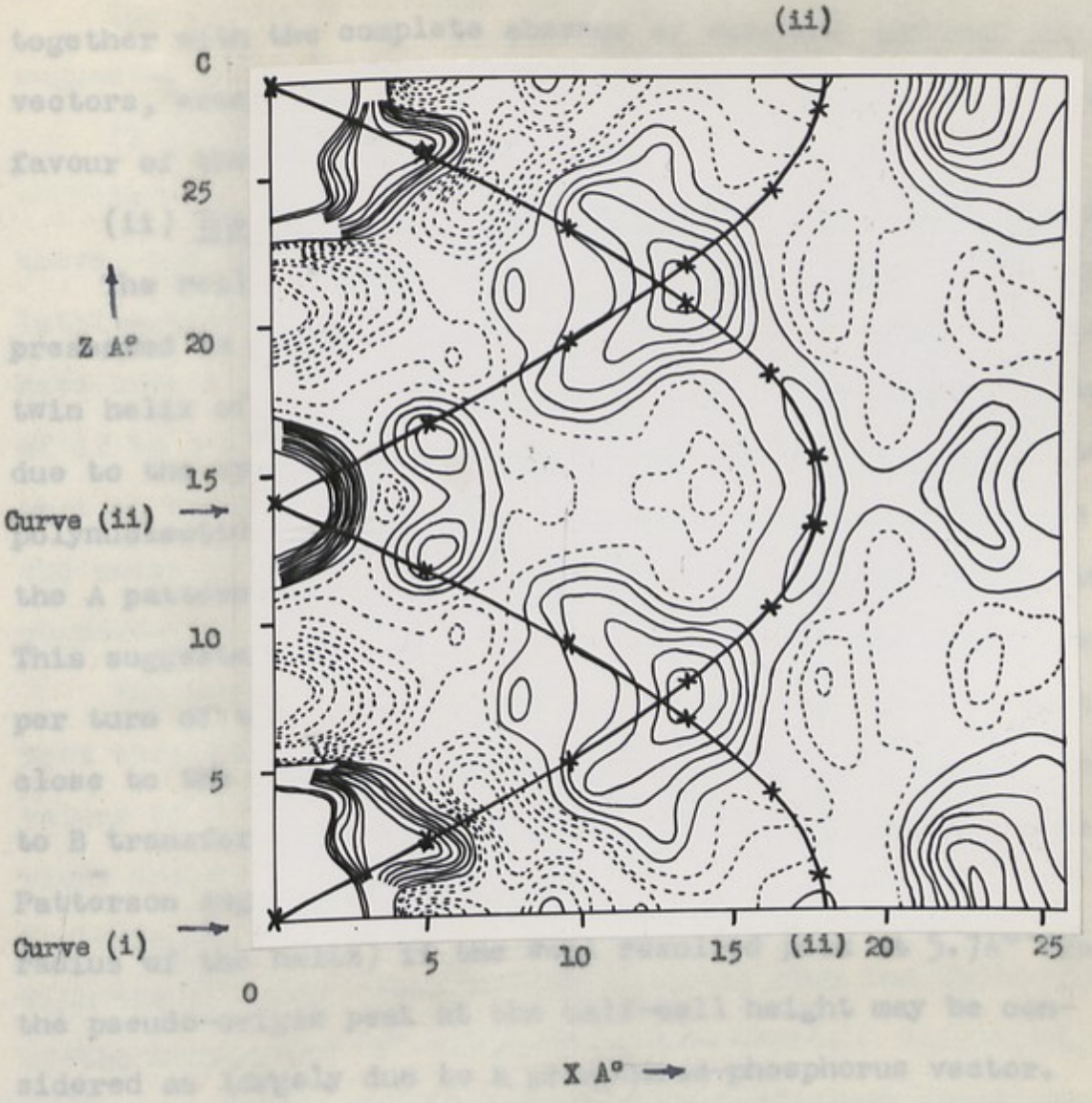


FIG. 5 Cylindrical Patterson for Structure A as shown in Fig. 4.

Curves (i) and (ii) are the Patterson functions of two co-axial "continuous line" helices of pitch equal to C, separation $C/2$ and radius 9Å.

X ≡ Theoretical peak for an intra-helical phosphorus-phosphorus vector, assuming 11 phosphorus atoms to lie on one turn of each of the helices represented by curves (i) and (ii)

general form of the positive regions of the Patterson. This, together with the complete absence of vertical and near vertical vectors, except at the half-cell height, is very strongly in favour of the helical model.

(ii) The number of phosphate residues per turn of the helix.

The real structure is not continuous. From the discussion presented in previous chapters it was to be expected that the twin helix of radius 9A°, suggested here for Structure A, was due to the spatial arrangement of the phosphate groups of two polynucleotide chains. Now the only meridional reflection in the A pattern is a rather weak one on the eleventh layer line. This suggested that there was possibly eleven phosphate residues per turn of the helix. This was reasonable, since it must be close to the number 10 found for the B structure because the A to B transformation is easily and rapidly reversible. The Patterson suggests this "elevening" (and also confirms the radius of the helix) if the well resolved peak at 5.7A° from the pseudo-origin peak at the half-cell height may be considered as largely due to a phosphorus-phosphorus vector.

For 5.7A° is just that distance between neighbouring phosphorus atoms if eleven are equally spaced along one turn of a helix of radius 9A°. Therefore in Fig.5, there are superimposed on the continuous helix curves, crosses, denoting the theoretical peaks of intra-helical phosphorus-phosphorus vectors, calculated on the basis of the following model; - two co-axial helices 9A° radius, pitch 28.1A°, with eleven equally spaced phosphorus atoms along each turn of each helix, the phosphorus atoms on one helix being directly above, in the

z direction, those in the other helix.

The agreement between the observed Patterson peaks and the suggested model is seen to be good.

(iii) The Unit Cell.

As well as giving structural information, as discussed above, the cylindrical Patterson also revealed the principle lattice vectors. The first region of high density that might have been a possible lattice translation occured at x values of 12 to 14A° and z between 5 and 9A°. It was found, however, that it was impossible to index the equatorial reflections on the basis of a unit cell having one parameter with an x-component of about 13A°.

The peaks around x = 22A°, z = 2A°; and x = 40A°, z = 0, were then considered as possible lattice translations. These values of x were close to the values of a and b used in previous index schemes from direct inspection of the X-ray diagram. Furthermore, the values a = 22A° and b = 40.0A° agreed well with the b/a sinβ ratio of 1.82 obtained from the application of the equatorial "d" values to a Bunn Chart. Slight adjustment around these values led to the satisfactory indexing of all the observed reflections on the basis of a c face-centered monoclinic cell having the parameters:-

a = 22.0A°

b = 39.8A° and β = 96.5°

c = 28.1A°

The agreement between the calculated and observed values of ξ was in most cases better than 1% and in no case worse than 2%. The values are listed in Table 1 Chapter VII.

(iv) The number of Polynucleotide chains, associated with each Lattice Point.

Due to the relatively small number of observed reflections and the ambiguity of indexing at large θ values, systematic absences could not be detected with certainty. However, if the unit cell is really c-face-centered monoclinic, mirror planes of symmetry are excluded by the presence of the asymmetric carbon atoms of the sugar rings, so that the space group can only be C_2. (ie diad axis along b).

If the phosphate "backbone" sequence of linkages really runs $C_5^1 - C_3^1 - C_5^1 - C_3^1$ through the sugar rings, then the "backbone" chain i.e. the strands of the 9A° radius helix, are non-centric. Therefore, the symmetry axis cannot pass through such a chain but must relate the chains together in pairs. Thus there must be an even number of chains associated with each lattice point.

The density and water up-take measurements on specimens at 75% R.H. have been discussed in Chapter IV. If a fibre water up-take of about 43% and a pellet density of 1.471 g/cc can be associated with Structure A, there are 22.7 nucleotides per lattice point. However, the density of the crystallites may be lower than 1.471 g/cc. This would reduce the value derived for the number of nucleotides per lattice point. Therefore the value of 22.7 is in fairly good accord with the concept of a system of two polynucleotide chains associated with each lattice point and having 11 nucleotides on each chain in the height of the unit cell. Since these two chains must be related by a diad axis they must run in opposite directions

with respect to the fibre axis (i.e. helical axis) as suggested by Watson and Crick.[6]

(v) Inter-Helical-Phosphorus-Phosphorus Vectors.

Assuming the helical model discussed above, the expected inter-helical P-P vectors were calculated, graphically, for a two strand helix centered on each lattice point. The orientation of the helix was determined from the three-dimensional Patterson and will be discussed in Chapter VIII.

The theoretical vector peaks calculated in this way are denoted by ⊖ in Fig. 6. It should be noted that each circle, i.e. a single inter-helical P-P vector, has only $1/11^{th}$ of the weight of the crosses shown in Fig.5 and 1/22 and 1/44th of the weight of the pseudo-lattice points (relating phosphate residues only) shown by crosses near the $z = \frac{1}{2}c$ level in Fig.6. The lattice points relating the whole Structure are denoted by larger crosses near to the origin level in C. Further, for correlation with the observed Patterson on which these theoretical peaks are placed, the weight of each peak should be considered as inversely proportional to x since in obtaining its cylindrical average it must be spread over a circle of radius x. If this is done the agreement between the suggested model and observation is reasonably good.

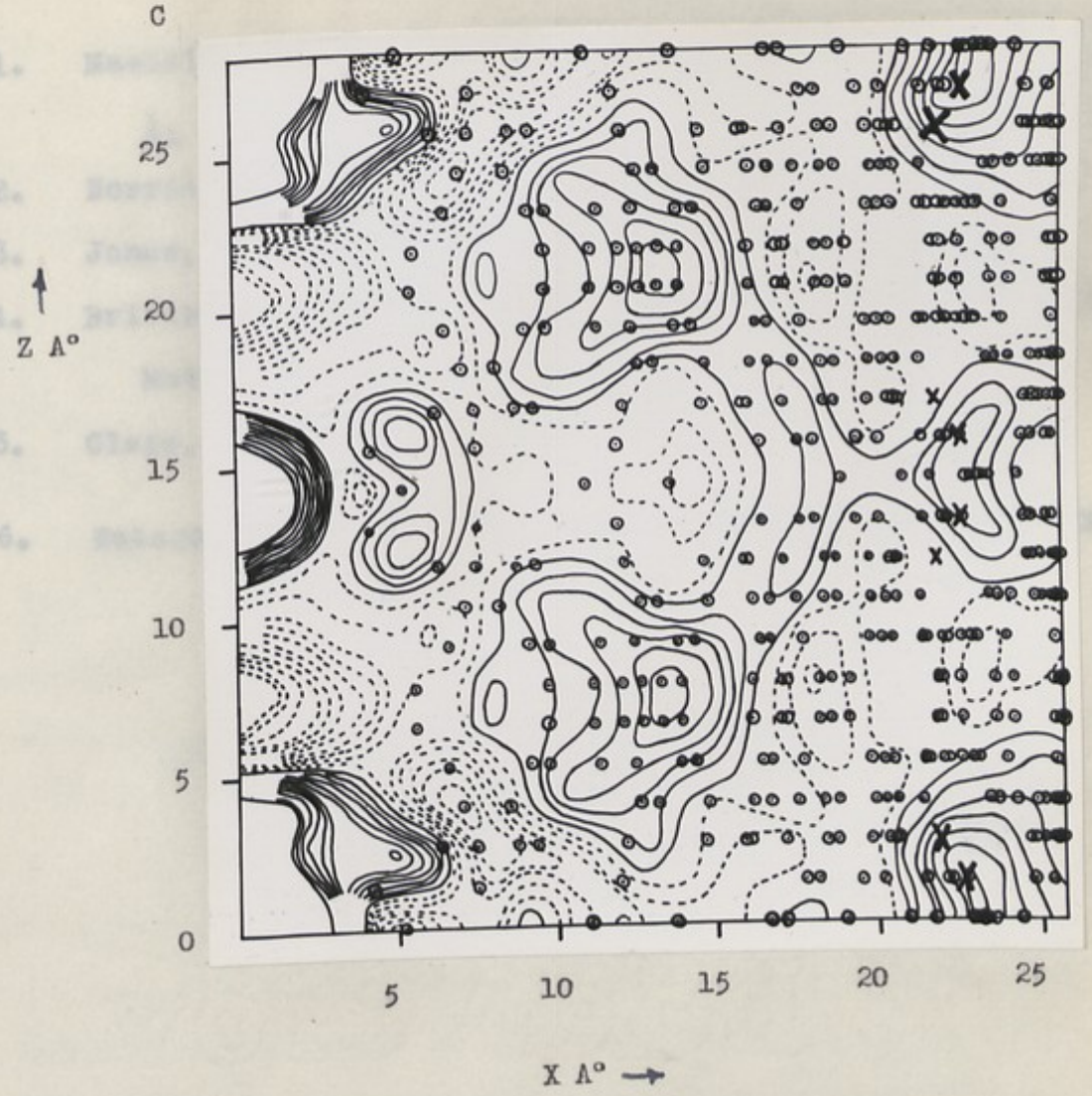


FIG. ~~PLATE~~ 6. Cylindrical Patterson function of Structure A

⊙ ≡ Theoretical peak for an inter-helical phosphorus-phosphorus vector.

REFERENCES IN CHAPTER VI.

1. MacGillavry, C.H. and Bruins, E.M. (1948) Acta. Cryst. 1. 156.

2. Norris, K.P. (1953) Ph.D. Thesis. London.

3. Jones, F.W. (1938) Proc. Roy. Soc. A. 166. 16.

4. British Association for the Advancement of Science. (1937): Mathematical Tables. 6. Part 1. C.U.P.

5. Clapp, M.M. (1937) J. Math. Phys. 16. 76.

6. Watson, J.D. and Crick, F.H.C. (1953) Nature. 171. 737.

CHAPTER VII

THE THREE DIMENSIONAL PATTERSON FUNCTION OF STRUCTURE A.

Indexing and $|F_{hkl}|^2$ Values.

The reflections observed in the X-ray diagram of Structure A were all indexed on the basis of the C-face-centered monoclinic unit cell discussed in the previous Chapter. In Table 1 the ξ values for the suggested indices, calculated on the basis of that unit cell, are compared with the observed values. Also listed in Table 1 are the intensity values, which have been corrected, as described in Chapter VI, and multiplied by an artifical temperature factor $e^{-(4.56)^2 s^2}$ ($s = 2 \sin \theta/\lambda$).

The correctness of the indexing scheme set out in Table 1 was confirmed, to some extent, by the double orientation photograph shown in Plate 6 Chapter IV. It may be seen, from that print, that many pairs of equivalent reflections in adjacent quadrants were of markedly different intensities. A list was drawn up in which those reflections which were strongest in the top left-hand quadrant of this diagram were labelled L, and those strongest in the top right-hand quadrant were labelled R. It was then found that all L reflections had been allotted indices hkl whereas all R reflections had indices $h\bar{k}l$. The distribution of the observed reflections among the different quadrants in R-space, as determined by the indexing suggested in Table 1, was thus directly confirmed experimentally.

It may be noticed, from Table 1, that for the larger values of θ no reflection could be indexed unambiguously. The intensities allotted to the various possibilities in such

TABLE I

DATA USED IN CALCULATION OF THE THREE DIMENSIONAL PATTERSON

$\frac{\sin\theta}{\lambda}.10^2$ OBS.	$\frac{\sin\theta}{\lambda}.10^2$ CALC.	hkl	$I_c.e^{-a^2s^2} = 2\lvert F_{hkl}\rvert^2$
ZERO LAYER			
5.25	5.21	110	9
8.80	8.81	130	104
9.00	9.14	200	30
10.00	10.04	040	3
10.50	10.42	220	4
13.25	13.35	150	13
13.8	13.57	240	5.5
	13.94	310	5.5
15.75	15.64	330	5
17.55	17.62	260	9
18.20	18.16	170	12
	18.28	400	9
18.85	18.72	350	4
	18.96	420	8
20.00	20.08	080	7
20.80	20.86	440	5
22.10	22.07	280	6
	22.29	370	6
22.85	22.99	510	10.5
	23.04	190	10.5
24.00	25.34	530	14
1st LAYER			$\lvert F_{hkl}\rvert^2$
4.90	4.86	$11\bar{1}$	22
5.50	5.57	111	11
8.60	8.61	$13\bar{1}$	39
9.05	9.02	131	23
10.10	10.08	$22\bar{1}$	10
13.10	13.22	$15\bar{1}$	16
15.30	15.29	$33\bar{1}$	7
16.10	16.00	331	7
22.60	22.59	$51\bar{1}$	6
	22.54	$37\bar{1}$	4
23.50	23.38	$46\bar{1}$	4
	23.67	$53\bar{1}$	5
2nd LAYER			
4.55	4.53	$11\bar{2}$	52
4.90	5.08	022	32
5.90	5.93	122	29
9.80	9.73	$22\bar{2}$	20
11.15	11.14	222	19
13.00	13.05	$24\bar{2}$	21
	13.14	$31\bar{2}$	21
17.20	17.21	$26\bar{2}$	31
18.05	18.00	$35\bar{2}$	9
	17.97	$17\bar{2}$	9
	18.05	262	9
20.00	20.10	082	2.5
	20.16	$44\bar{2}$	2.5
21.9	21.80	$37\bar{2}$	6
	21.75	$28\bar{2}$	6
22.90	22.79	$37\bar{2}$	5
	22.90	$19\bar{2}$	5
	23.08	$46\bar{2}$	5
26.60	26.63	$48\bar{2}$	1
	26.86	$55\bar{2}$	1
	26.45	$2.10.\bar{2}$	1
	26.62	$60\bar{2}$	1
27.90	27.63	$1.11.\bar{2}$	1.5
	28.20	$57\bar{2}$	1.5
3rd LAYER			
12.80	12.80	$24\bar{3}$	3.5
	12.76	$31\bar{3}$	3.5
14.55	14.41	$24\bar{3}$	4
	14.60	$33\bar{3}$	3
14.95	15.11	063	4
	15.12	$31\bar{3}$	4
19.80	19.82	$44\bar{3}$	2
21.50	21.60	$28\bar{3}$	2
4th LAYER			
12.35	12.37	$31\bar{4}$	1.5
	12.56	$22\bar{4}$	1.5
14.10	14.16	$15\bar{4}$	10
	14.26	$33\bar{4}$	9

TABLE I. contd.

$\frac{\sin\theta}{\lambda}.10^2$ OBS.	$\frac{\sin\theta}{\lambda}.10^2$ CALC.	hkl	$I_c.e^{-a^2s^2} = \lvert F_{hkl}\rvert^2$
4th LAYER contd.			
14.95	14.70	22$\bar{4}$	13.5
	15.15	06$\bar{4}$	13.5
16.90	16.84	26$\bar{4}$	14
	17.06	33$\bar{4}$	4
21.45	21.34	37$\bar{4}$	2.5
	21.43	28$\bar{4}$	2.5
22.75	22.77	28$\bar{4}$	2
	22.78	19$\bar{4}$	3
5th LAYER			
11.90	11.98	31$\bar{5}$	12
12.40	12.32	24$\bar{5}$	6
	12.22	22$\bar{5}$	6
14.0	12.34	15$\bar{5}$	3
6th LAYER			
8.40	8.40	22$\bar{6}$	12
10.4	10.33	04$\bar{6}$	14
	10.26	13$\bar{6}$	14
11.70	11.59	31$\bar{6}$	12
12.75	12.93	15$\bar{6}$	12
13.50	13.58	33$\bar{6}$	21
7th LAYER			
7.90	7.79	11$\bar{7}$	4
	7.74	13$\bar{7}$	5
11.0	11.19	31$\bar{7}$	11
12.1	11.94	20$\bar{7}$	5
	11.87	24$\bar{7}$	5
13.2	13.26	33$\bar{7}$	4
	12.95	22$\bar{7}$	6
8th LAYER			
10.80	10.81	31$\bar{8}$	10
	10.82	13$\bar{8}$	6
12.50	12.83	15$\bar{8}$	9
	12.34	20$\bar{8}$	6
15.70	15.89	42$\bar{8}$	3
9th LAYER			
5.7	5.54	20$\bar{9}$	2

NB I_c = Observed Intensity corrected as described in Chapter VI.

$a = 4.56.$

$s = 2 \sin\theta/\lambda$.

cases were obtained by arbitrary division of the corrected intensity value for the observed, multiple, reflection. In cases where the possible reflections included hkl and $\bar{h}\bar{k}l$ the double orientation photograph was taken as a guide to the weighting used in this division of intensity. If the ambiguous reflection in question was strong in the R quadrant (as defined above), then the $\bar{h}\bar{k}l$ alternative was weighted strongly. If the observed reflection was strongest in the L quadrant then the possible hkl value was given most weight.

Those reflections at high angles which should have been well resolved and single were not observed. This seemed to imply that it was only where the geometry of the reciprocal lattice was such that two or more reflections reinforced one another, that a photographic effect was observed against the rather strong diffuse background. Therefore in any calculations involving the use of the indices of the reflections an artificial temperature factor was doubly important; it was necessary to minimise errors due to the unavoidable ambiguity of indexing the far out reflections, it was also necessary to minimise "cut off" effects due to the fact that the corrected intensity values did not decrease appreciably with increasing θ.

<u>The Patterson Function for Space Group C_2.</u>

The vector density at any point xyz in the unit cell is given by the Patterson Function:-

$$P(xyz) = \sum_h \sum_k \sum_{l=-\infty}^{+\infty} |F_{hkl}|^2 \cos 2\pi(hx + ky + lz) \qquad (1)$$

In the monoclinic cell found for Structure A, space group C2,

the following reflections are equivalent:-

$hkl = \bar{h}k\bar{l} = h\bar{k}l = \bar{h}\bar{k}\bar{l}$ $\quad hkl \neq \bar{h}kl$

and $\bar{h}kl = hk\bar{l} = \bar{h}\bar{k}l = h\bar{k}\bar{l}$

Therefore, equation (1) may be reduced to the form

$$P_{(xyz)} = \sum_h \sum_k \sum_0^\infty (|F_{\bar{h}kl}|^2 + |F_{hkl}|^2) \cdot \cos 2\pi hx \cdot \cos 2\pi ky \cdot \cos 2\pi lz.$$

$$+ \sum_h \sum_k \sum_0^\infty (|F_{\bar{h}kl}|^2 - |F_{hkl}|^2) \cdot \sin 2\pi hx \cdot \cos 2\pi ky \cdot \sin 2\pi lz.$$

Multiplicity

In applying equation (2) the multiplicity of the $|F_{hkl}|^2$ ($h \pm$) terms should be considered. However, since the intensity data for Structure A was obtained from a complete rotation diagram, no multiplicity corrections were required for values obtained from layer lines with l greater than zero. It only remained to halve the observed equatorial and meridional intensities. The values of the F^2 terms used in evaluating equation (2) are listed in Table 1.

Calculation.

The summations of equation (2) were carried out in four stages, using Beevers Lipson strips. The relative field figures were calculated at an interval of $a/30$ in x, $b/60$ in y and $c/30$ in z, for a quarter cell.

Interpretation of the Patterson Function.

The resulting three dimensional Patterson function is shown in Fig. 1, (i) and (ii), as a series of sixteen a - b quarter cell sections at intervals in z of $c/30$, from $z = 0$ to $z = c/2$. The contours were drawn at equal intervals, positive values being indicated by a full line and negative values by a dotted line.

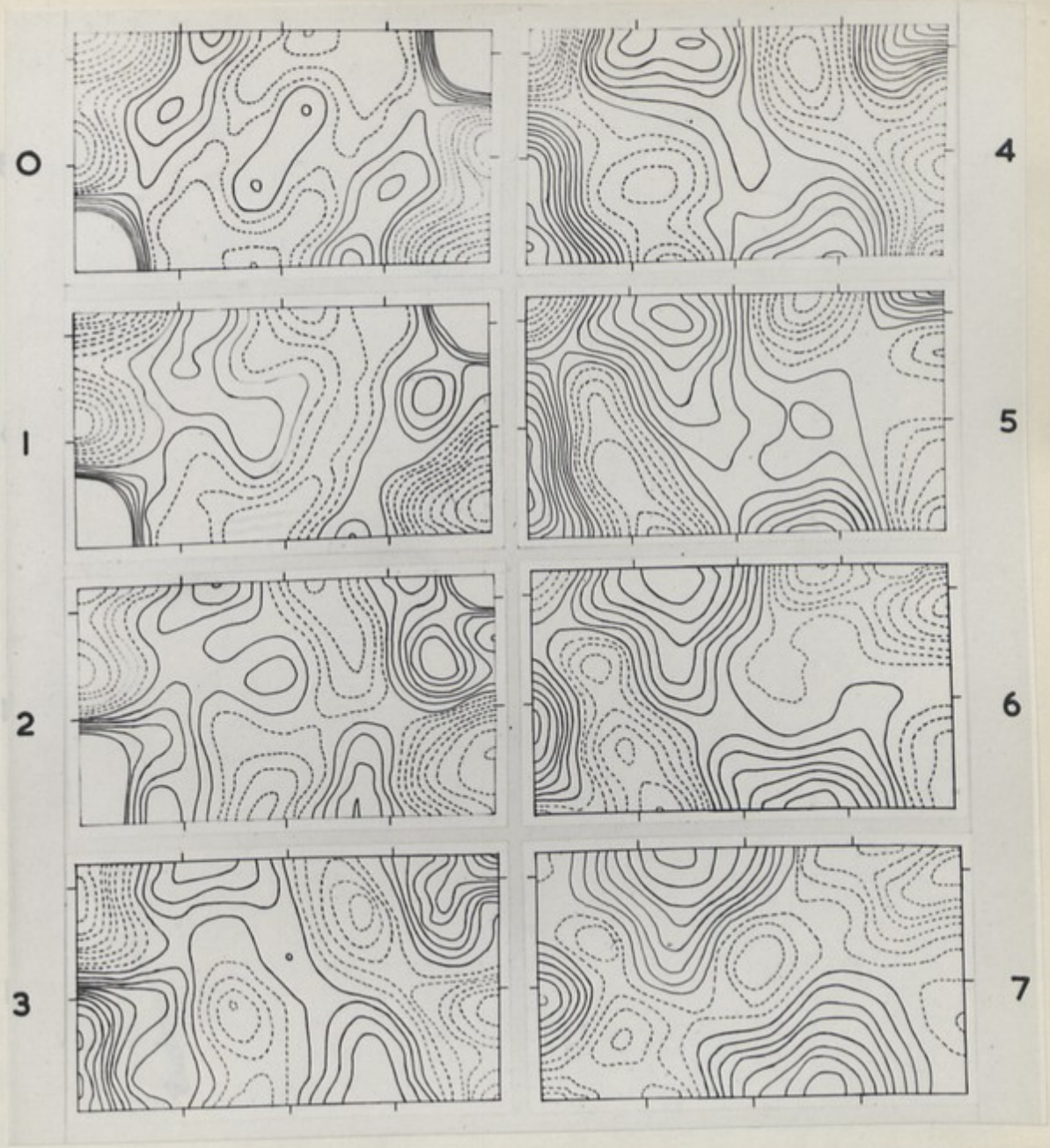


FIG. 1 (i) THE THREE DIMENSIONAL PATTERSON FUNCTION OF STRUCTURE A

Eight a-b quarter sections at intervals in z of $^{C}/_{30}$ from z = 0 to z = $^{7}/_{30}$.C, contours are drawn of the relative field figures at equal intervals, positive values being denoted by a full line and negative values by a dotted line. The scale markings along the edges of the sections correspond to an interval of 5A°.

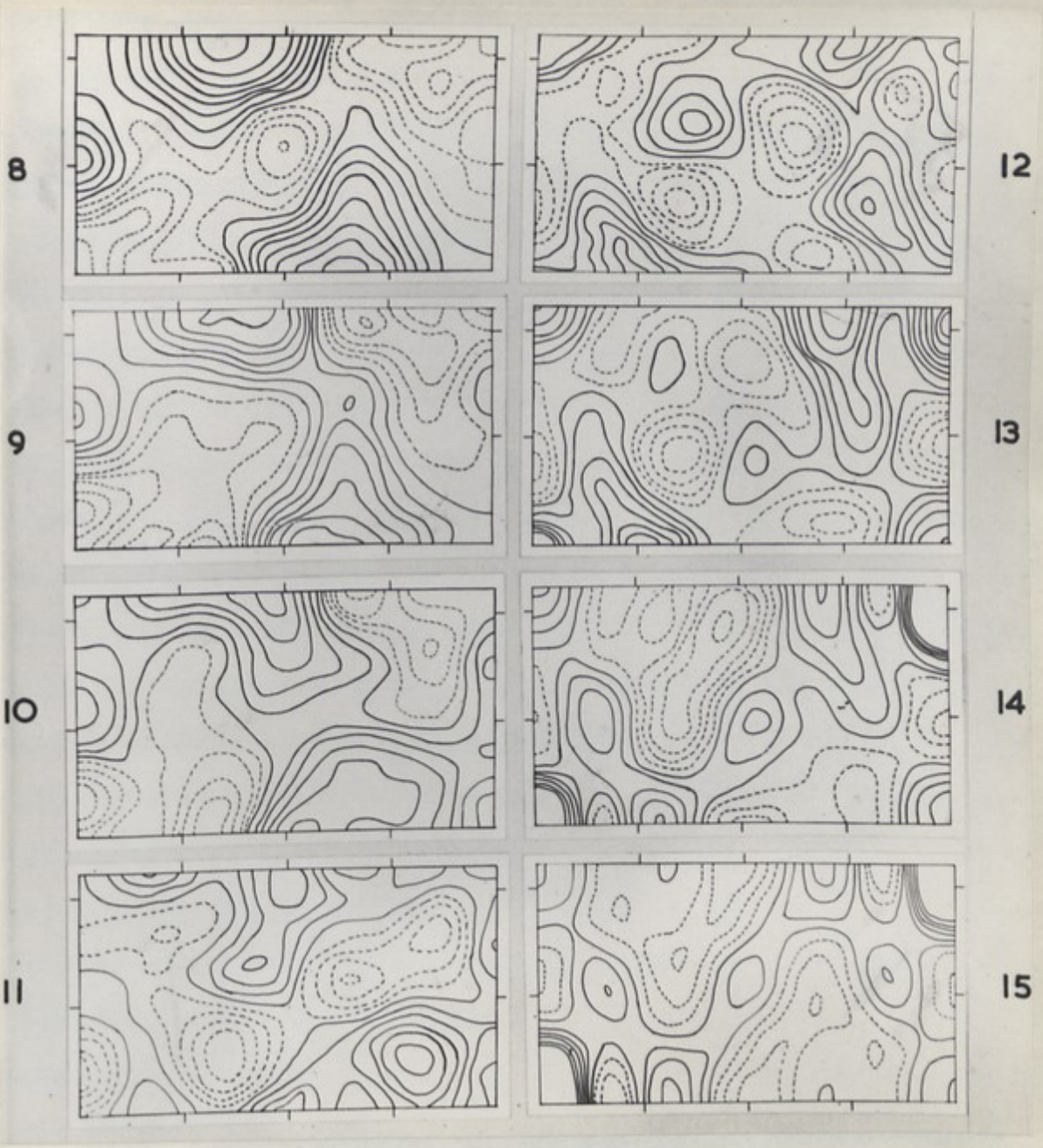


FIG. 1. (ii) THE THREE DIMENSIONAL PATTERSON FUNCTION OF STRUCTURE A

Eight a-b quarter sections at intervals in z of $C/30$ from $z = 7/30\ C$ to $z = C/2$; contours are drawn of the relative field figures at equal intervals, positive values being denoted by a full line and negative values by a dotted line. The scale markings along the edges of the sections correspond to an interval of 5A°.

The very strong peak at the half cell height, x = y = 0 z = 14.05A°, was expected from the cylindrical Patterson. However, many features of the zero section other than the origin peak are almost entirely reversed in the half-cell section (15). This tends to confirm the suggestion, made in Chapter VI, that the phosphate groups of Structure A repeat in the period $z = c/_2$ but, since the molecule consists of two chains related by a diad axis, the rest of the molecule only repeats in the full cell period. Therefore the a - b sections were examined for agreement with the model suggested from a study of the cylindrical Patterson, discussed in Chapter VI. That is, a 9A° radius two strand helix of phosphorus atoms, the strands being separated by $z = c/_2$ and the eleven equally spaced phosphorus atoms in one strand being exactly above those in the other strand, (in the z direction). (Pitch = 28·1A°).

(i) Orientation of the Helix.

In order to calculate the co-ordinates of the vector peaks to be expected from the proposed model it was first necessary to decide on the orientation of the helix in the unit cell. Consideration of the density had shown, Chapter VI, that there was only one helical unit per primitive unit cell. Therefore each helix could be considered as centered on a lattice point. If the helix could be right-handed or left-handed, there were four possible arrangements of the phosphorus atoms about a lattice point that could comply with the symmetry of the space group. These are shown in Fig.2.

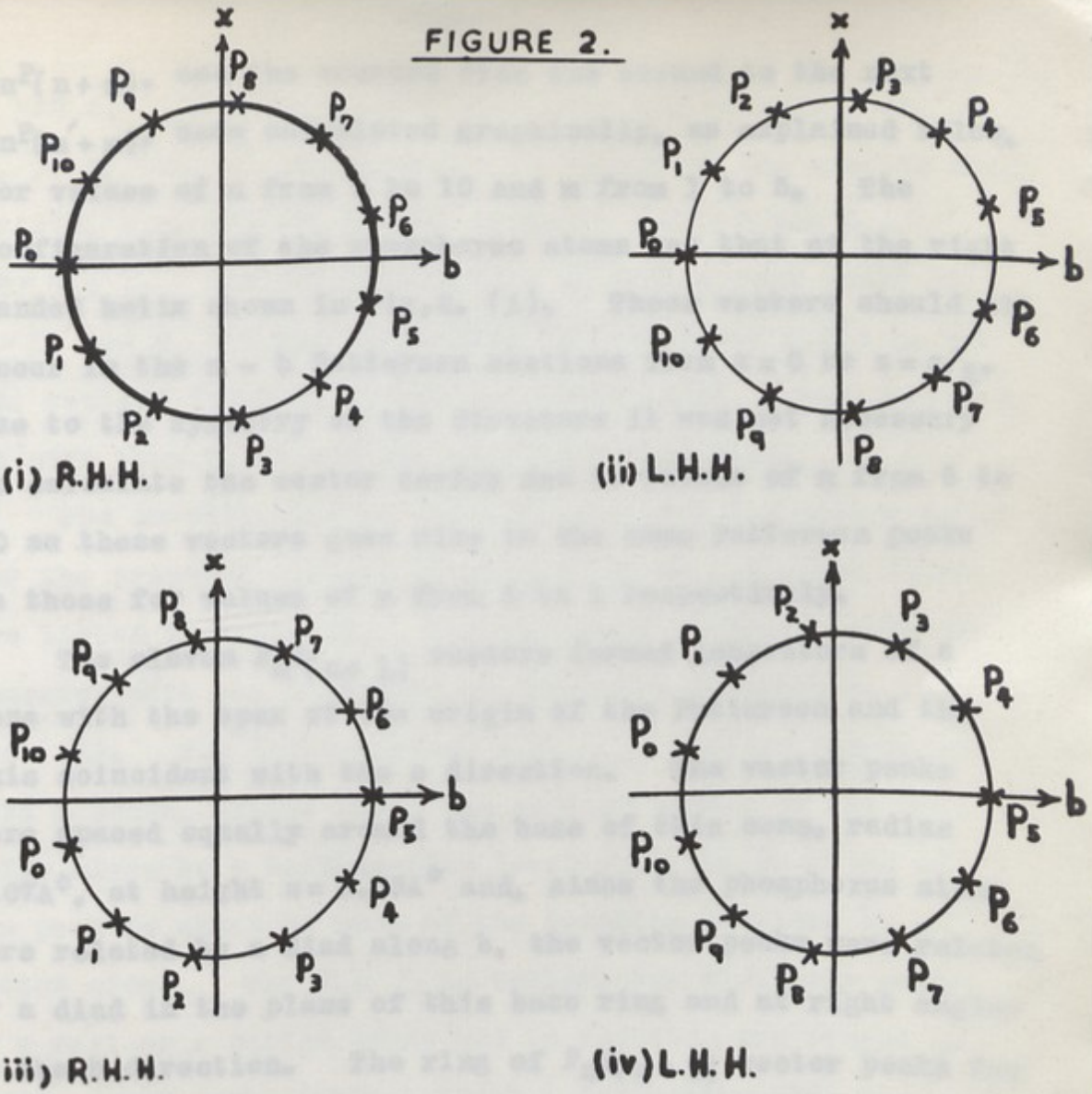
as projections of a turn of one strand of the helix onto a plane at right angles to the c axis through the level of z = 0. Since the phosphate-sugar backbone chain is non-centric the symmetry axes along b cannot pass through a chain, but must relate one chain, n, to the other, n′. Thus in arrangements (i) and (ii) P_0 must be placed at the quarter cell height and so related by the diad axis at c/2 to the atom $P_{0'}$ at z = 3/4 c. In arrangements (iii) and (iv) P_5 must be considered to be at the quarter cell height. With regard to the P-P vectors that were to be expected from these four possible configurations, it may be seen from Fig.2. that the vectors from (i) and (iv) are the same, as also are those from (ii) and (iii). Hence from the Patterson function alone it was not possible to decide whether the helix was right-handed or left-handed, but only between two possible pairs of configurations.

Watson and Crick have suggested[1], from model building, that a helical configuration of DNA could only be right-handed. Adopting that suggestion the possible arrangements shown in Fig.2. are reduced to (i) and (iii). In (i) there is one of the eleven $P_n P_{(n+1)}$ vectors in the positive "a" direction with zero b component, i.e. P_5P_6 which has a = 5.07A° and z = 2.55A°. In (iii) the $P_nP_{(n+1)}$ vector with zero b component is $P_{10}P_{0}$ and is in the negative "a" direction. Inspection of the Patterson sections 2 and 3 showed that only arrangement (i) was possible.

(ii) Intra-Helical Vectors. $P_nP_{(n+m)}$ and $P_nP_{(n'+m)}$.

The intra-helical phosphorus vectors for one strand,

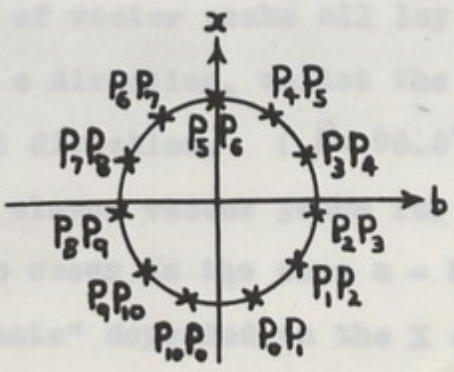
FIGURE 2.



PROJECTIONS OF THE FOUR POSSIBLE HELICAL ARRANGEMENTS OF THE PHOSPHORUS ATOMS IN ONE TURN OF STRUCTURE A.

R.H.H. = RIGHT-HANDED HELIX
L.H.H. = LEFT-HANDED HELIX

FIGURE 3.



THE CONFIGURATION OF THE 11 P_0P_1 TYPE VECTOR PEAKS FROM ARRANGEMENT (i) SHOWN IN FIGURE 2.

$P_nP_{(n+m)}$, and the vectors from one strand to the next $P_nP_{(n'+m)}$, were calculated graphically, as explained below, for values of n from 0 to 10 and m from 1 to 5. The configuration of the phosphorus atoms was that of the right handed helix shown in Fig.2. (i). These vectors should all occur in the a - b Patterson sections from $z = 0$ to $z = c/2$. Due to the symmetry of the Structure it was not necessary to calculate the vector series due to values of m from 6 to 10 as these vectors gave rise to the same Patterson peaks as those for values of m from 5 to 1 respectively.

The eleven $P_nP_{(n+1)}$ vectors formed generators of a cone with the apex at the origin of the Patterson and the axis coincident with the c direction. The vector peaks were spaced equally around the base of this cone, radius 5.07A°, at height $z = 2.55$A° and, since the phosphorus atoms were related by a diad along b, the vector peaks were related by a diad in the plane of this base ring and at right angles to the b direction. The ring of $P_nP_{(n+1)}$ vector peaks for the arrangement of phosphorus atoms shown in Fig.2 (i) is reproduced in Fig.3. The rest of the vector peaks mentioned above, i.e. m = 2 to 5, were drawn out graphically in this way. However, these rings of vector peaks all lay in planes perpendicular to the c direction, whilst the a - b sections were inclined to that direction. ($\beta = 96.5°$). Therefore not all the proposed eleven vector peaks for any one value of m were expected to occur in the same a - b section. The "section co-ordinate" depended on the X co-ordinate of the vector peak, the X axis being mutually perpendicular to

b and c. If a proposed vector peak had co-ordinates X, y, z, then it was expected to appear on the section whose origin, i.e. a=b = 0, had co-ordinate $z' = z + X.\tan\theta$., ($\theta = 6\frac{1}{2}^{o}$). The "a" co-ordinate of such a peak would be a= X. sec θ, but the difference between X and X.sec θ, for the vectors involved in the suggested model, never amounted to more than 0.1A^{o} and was therefore neglected.

The co-ordinates of the intra-helical P-P vector peaks for the proposed structure, calculated as described above, are listed in Table II. The positions of these peaks are compared with the Patterson of Structure A in Fig.4. (i) and (ii). The intra-helical $P_nP_{(n+m)}$ vectors are marked with a cross, except where the section number of the peak falls close to a half-section level when the position of the peak is shown on each of the bounding sections by a dotted cross. The $P_nP_{(n'+m)}$ vectors are similarly marked by a full or a dotted cross and in addition are ringed by a full line. The <u>inter-helical</u> P.P. vectors give rise to vector peaks which in the Patterson are indentical with the intra-helical P.P. peaks, since there is one helix centered on each lattice point. It may be concluded from Fig.4. that the suggested helical array of phosphorus atoms is in good agreement with the three dimensional Patterson diagram of Structure A.

TABLE II

Intra-Helical Vectors $P_nP_{(n+m)}$ and $P_nP_{(n'+m)}$.

Due to the symmetry of the suggested Structure it was only necessary to calculate the positions of the vector peaks expected for values of m from 1 to 5. The peaks for any value of m, (n from 0 to 10), are related by a diad at right angles to the c - b plane, therefore only six of the possible eleven peaks are listed below. The co-ordinates were calculated graphically and are correct to ± 0.1A°. The X and y co-ordinates of the $P_nP_{(n'+m)}$ vector peaks are the same as those of the corresponding $P_nP_{(n+m)}$ peaks, whilst the correct section height is given by adding 15 to the appropriate $P_nP_{(n+m)}$ value.

$P_nP_{(n+m)}$	X A°	y A°	a - b Section	
			$P_nP_{(n+m)}$	$P_nP_{(n'+m)}$
$P_nP_{(n+1)}$	Z = 2.55A°:-			
P_0P_1	$\bar{4}$.85	1.35	2.10	17.10
P_1P_2	$\bar{3}$.35	3.75	2.30	17.30
P_2P_3	$\bar{0}$.75	4.95	2.60	17.60
P_3P_4	2.10	4.55	3.00	18.00
P_4P_5	4.25	2.70	3.25	18.25
P_5P_6	5.07	0	3.35	18.35
$P_nP_{(n+2)}$	Z = 5.10A°:-			
$P_{10}P_1$	$\bar{9}$.70	0	4.25	19.25
P_0P_2	$\bar{8}$.20	5.25	4.45	19.45
P_1P_3	$\bar{4}$.10	8.85	4.95	19.95
P_2P_4	1.35	9.65	5.60	20.60
P_3P_6	6.35	7.30	6.20	21.20
P_4P_6	9.30	2.80	6.60	21.60

TABLE II contd.

$P_nP_{(n+m)}$	X A°	y A°	a - b Section	
			$P_nP_{(n+m)}$	$P_nP_{(n'+m)}$
$P_nP_{(n+3)}$	Z = 7.65A°:-			
$P_{10}P_2$	$\overline{13}$.10	3.70	6.60	21.60
P_0P_3	$\overline{9}$.0	10.20	7.10	22.10
P_1P_4	2.10	13.50	7.90	22.90
P_2P_5	5.55	12.40	8.85	23.85
P_3P_6	11.45	7.35	9.55	24.55
P_4P_7	13.60	0	9.85	24.85
$P_nP_{(n+4)}$	Z = 10.20A°:-			
P_9P_2	$\overline{16}$.35	0	8.90	23.90
$P_{10}P_3$	$\overline{13}$.70	8.85	9.20	24.20
P_0P_4	$\overline{6}$.80	14.80	10.10	25.10
P_1P_5	2.30	16.10	11.15	26.15
P_2P_6	10.75	12.30	12.20	27.20
P_3P_7	16.25	4.75	12.85	27.85
$P_nP_{(n+5)}$	Z = 12.75A°:-			
P_9P_3	$\overline{17}$.15	4.85	11.50	26.50
$P_{10}P_4$	$\overline{11}$.85	13.30	12.20	27.20
P_0P_5	$\overline{2}$.70	17.55	13.30	28.30
P_1P_6	7.35	16.20	14.50	29.50
P_2P_7	15.00	9.60	15.45	30.45
P_3P_8	17.80	0	15.80	30.80

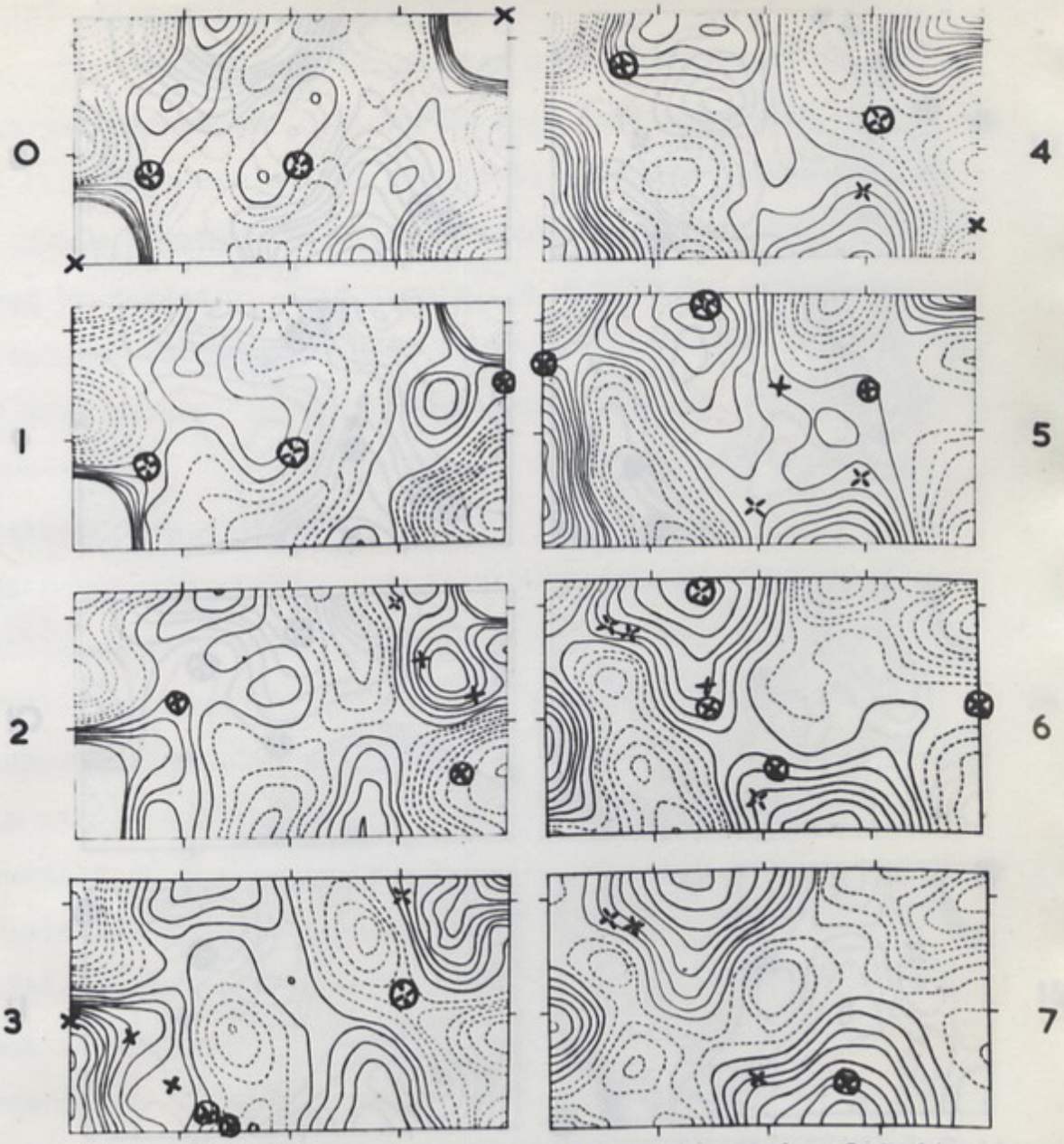


Fig. 4. (i) The comparison of the three dimensional Patterson function of Structure A, sections 0 to 7, with the P-P vectors that would be expected if the model outlined below was associated with each lattice point.

Model Structure: A helix consisting of two co-axial strands of 9A° radius and 28.1A° pitch (= c), with 11 phosphorus atoms spaced equally along each turn of each strand, the axis of the helix being coincident with the c direction. The separation of the strands is $\frac{1}{2}$.c and the corresponding successive phosphorus atoms in each strand, P_n and P_n', have the same co-ordinates in a plane at right angles to the axis of the helix.

X = $P_nP_{(n+m)}$ vector peak. ✕ = $P_nP_{(n+m)}$ vector peak lying close to a half-section level.

The full or broken cross enclosed by a circle denotes that the conditions defined above are applied to a $P_nP_{(n'+m)}$ vector peak.

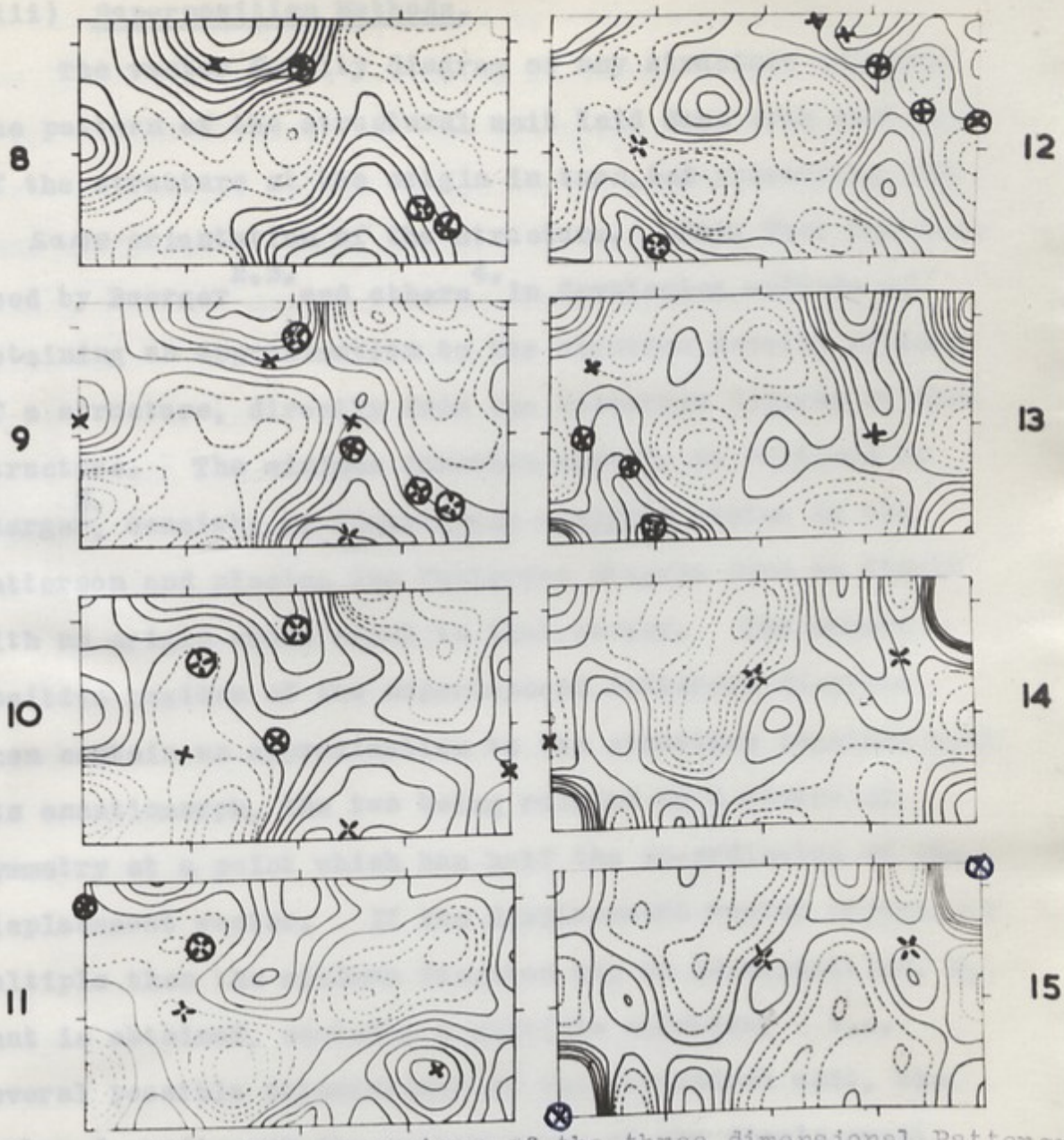


Fig. 4. (ii) The comparison of the three dimensional Patterson function of Structure A, a - b sections 8 to 15, with the P-P vectors that would be expected if the model defined in the caption to Fig.4. (i) was associated with each lattice point.

$\times = P_n P(n+m)$ vector peak.

$\times\!\!\!\!\diagup = P_n P(n+m)$ vector peak lying close to a half-section level.

The full or broken cross enclosed by a circle denotes that the conditions defined above are applied to a $P_n P(n'+m)$ vector peak.

(iii) Superposition Methods.

The vector density diagram of any structure contains the pattern of the structural unit laid down with each atom of the structure at the origin in turn, but presenting the same orientation of the structure. This fact has been used by Buerger[2.3.] and others[4.] in developing methods of obtaining an approximation to the electron density pattern of a structure, directly from the Patterson diagram of that structure. The minimum function method, as outlined by Buerger[3.], consists of identifying a single vector in the Patterson and placing the Patterson diagram down on itself with an origin shift equal to that vector. The common positive regions of the superimposed Patterson diagrams then contain an approximation to the structure together with its enantiomorph, the two being related by a centre of symmetry at a point which has half the co-ordinates of the displacement vector. If the displacement vector chosen was multiple then the minimum function map or superposition, M_2, that is obtained, contains a multiple solution; i.e. several possible orientations of the structural unit, the number depending on the multiplicity of the displacement vector, together with the respective enantiomorphs. Maps of the minimum function of higher rank, i.e. maps approximating more closely to the electron density, can be obtained by means of a step-wise procedure. Two M_2 maps are formed, based upon two different displacement vectors, these maps may then be combined provided that the spatial relationship of the two displacement vectors in the structure

is known. Each map must also be weighted according to the type of atoms producing the Patterson peak used as the displacement vector. If the two M_2 maps to be combined are derived from the use of two different Patterson peaks, but each is caused by the interactions of the same kind of atoms, then the maps may be combined directly without any relative scaling of their contours being required. In this way if two or three superpositions are made using vectors whose mutual relationship in the structure is known, these superpositions may then be combined to eliminate all but one orientation of the structural pattern.

An attempt was made to apply the superposition method, as outlined above, to the three dimensional Patterson function of Structure A without making any assumptions as to the structural pattern. It was realised that the resolution of the Patterson was too low to permit of a detailed solution, but, since the major contribution to the vector density was most likely to be vectors between the heavy phosphate groups, it was expected that it might be possible to find the approximate configuration of those groups in the Structure. Accordingly, small and where possible isolated, peaks in the a - b sections shown in Fig.1. (i) and (ii) were used as displacement vectors. It was thought that such peaks were most likely to be single vectors, or at the least to be vectors of low multiplicity. From the sixteen a - b sections of Fig. 1. sixteen superposition sections were obtained for each of the twelve displacement vectors taken. The origin of each set of superposition sections

was made to coincide with the centre of symmetry of the superposition function, which occurred at the centre of the displacement vector. In order to discover how to combine these various superpositions, to obtain functions of a higher rank, an attempt was made to identify a common vector in each superposition. The superpositions were then combined in pairs, so that the common vectors in each were coincident. The resultant of one pair of superpositions was then combined with a third superposition and the resulting pattern was examined for the presence of some continuous chain structure.

In all the superpositions that were drawn the presence of a diad axis in the b direction was strongly apparent, in no case was there any evidence of a mirror plane or a glide plane. This was taken as a strong indication that the method was giving a partially true solution, since the space group used in calculation of the Patterson was C_2. However, the strong vectors in the a - c plane which as may be seen from Fig.4 are not accounted for by any circularly symmetrical structure, always gave rise to apparent chains in the direction of the a - c diagonal repeating at an interval of one-third in b. These vectors may be seen more clearly in Fig. 5, (i) (ii) which shows the a - c sections for $y = b/3$ and $y = 0$. The direct application of the minimum function method therefore failed to show any evidence of a chain structure compatible with that suggested by the cylindrical Patterson. It may be concluded, from this failure, that the method is very sensitive to the exact positioning of the displacement vectors used, and to the correct weighting

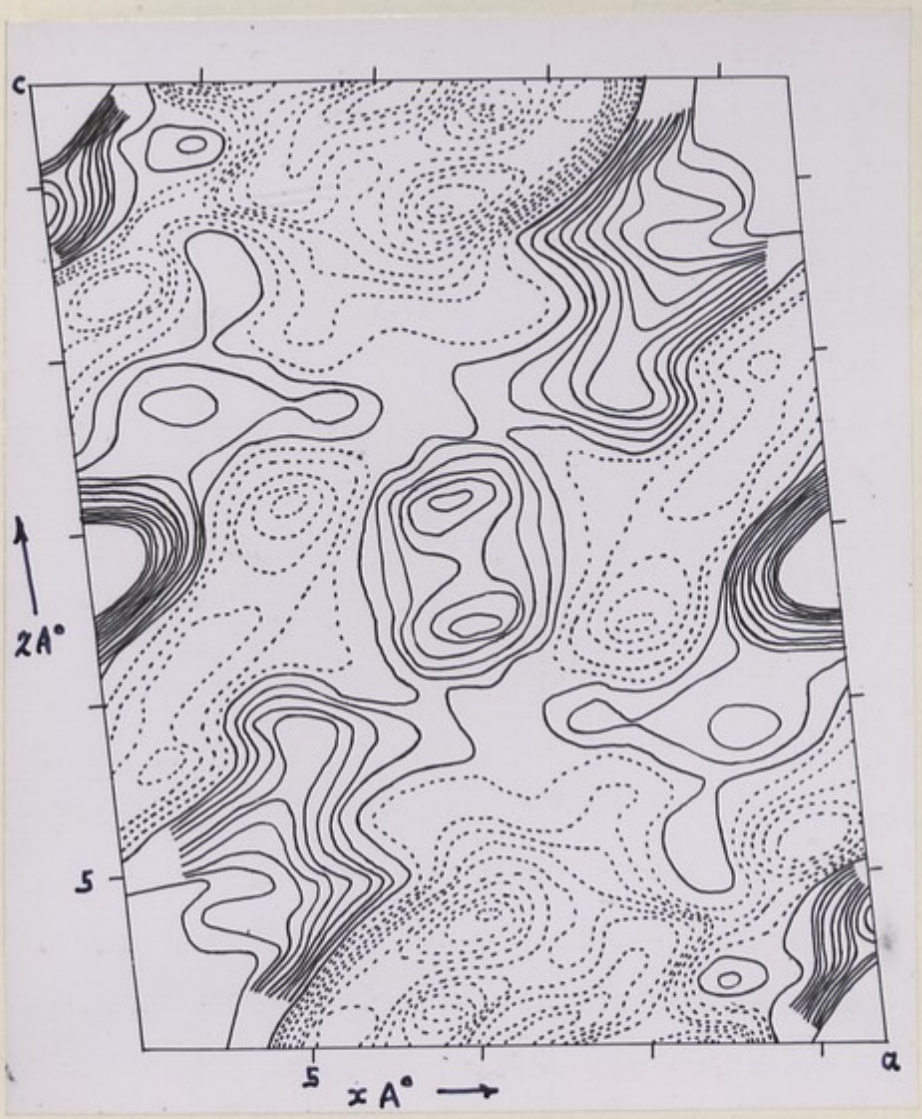


Fig. 5. (i) The a - c section y= 0, of the three dimensional Patterson function of Structure A.

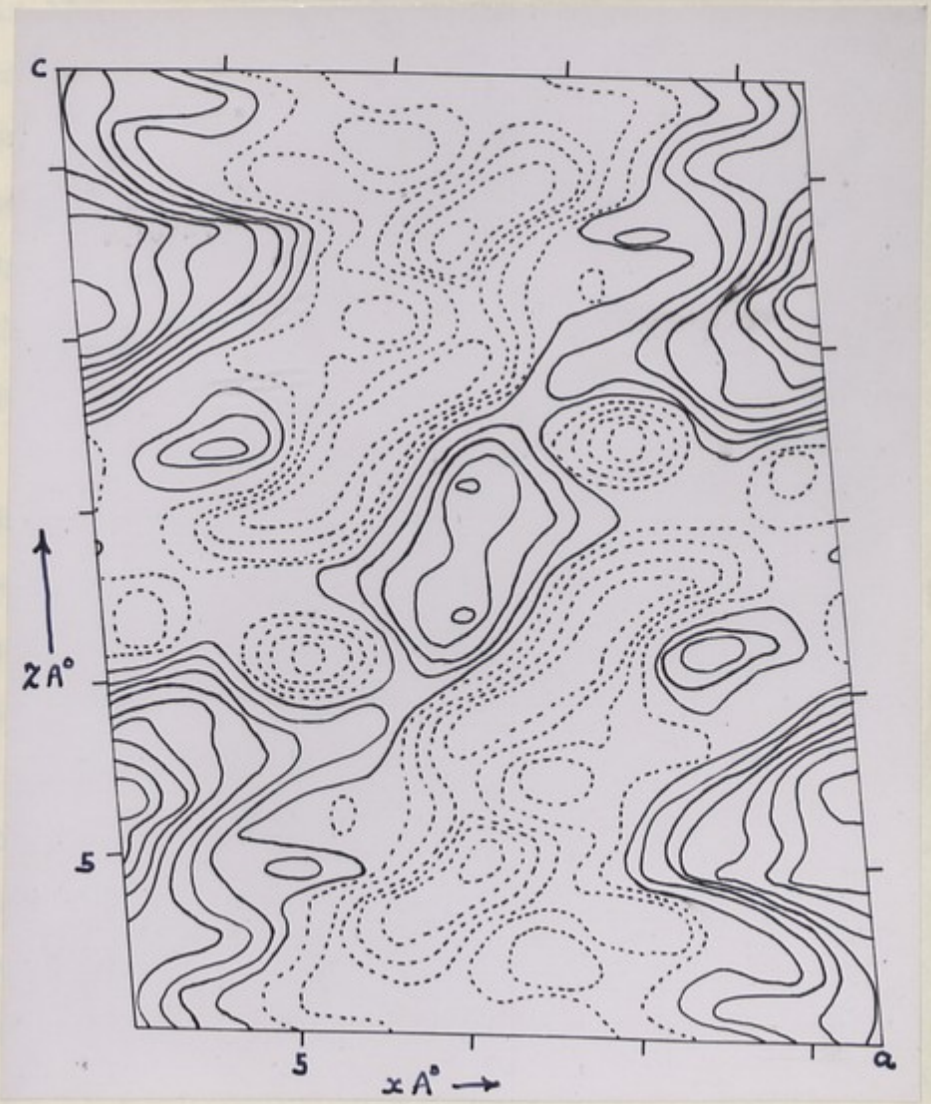
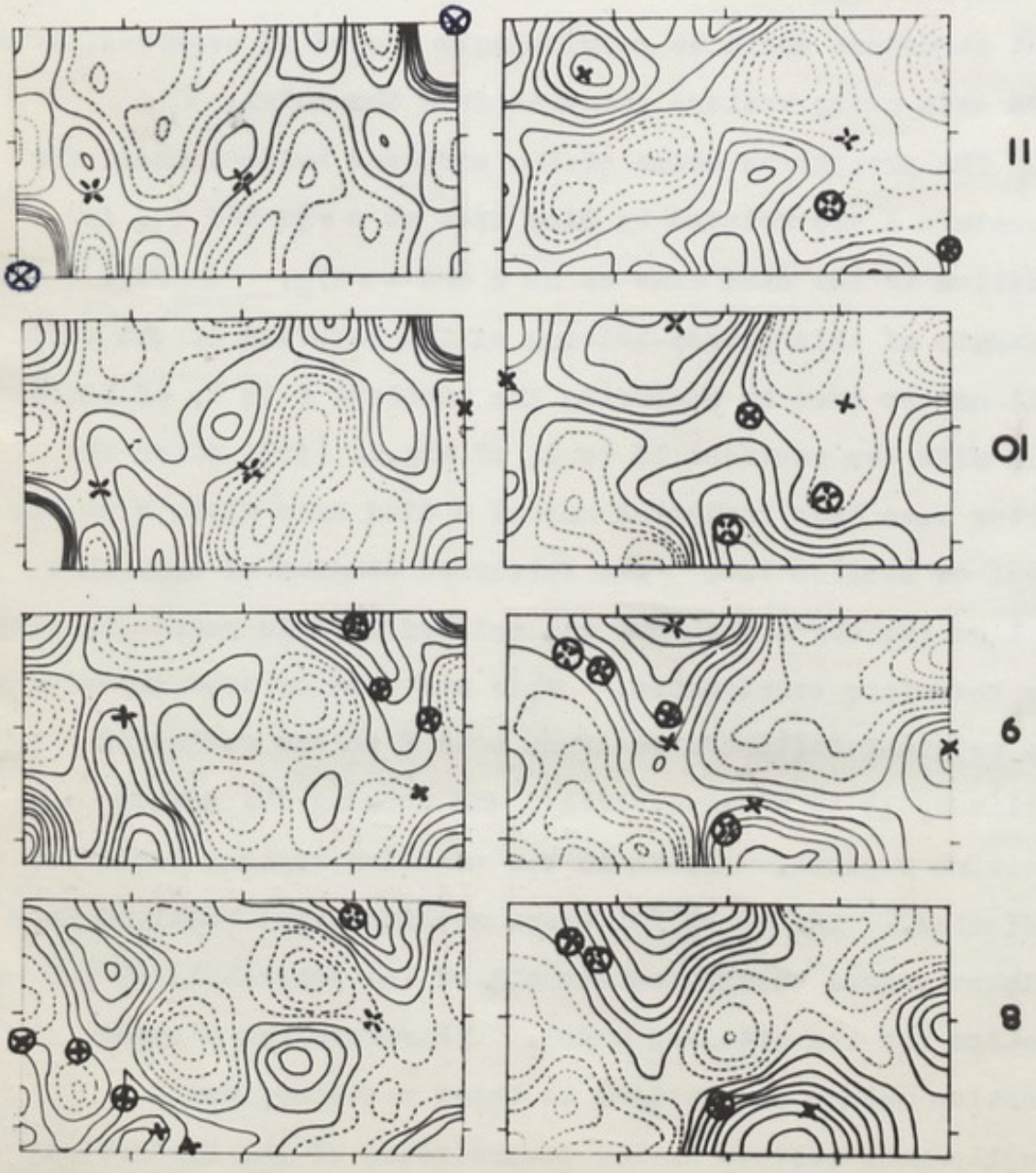


Fig. 5. (ii) The a - c section y = b/3, of the three dimensional Patterson function of Structure A.

of the superpositions when combining them to obtain higher order functions. Therefore it would seem that in considering the Patterson of any structure with a large number of atoms in the unit cell, unless some structural assumptions may be made which would permit the identification of vector peaks and the correct combination of superpositions, a structure solution by the minimum function method is probably impossible. By the same argument, if a model structure were available then that structure could be convincingly tested if superpositions were made using vectors expected from the model.

The phosphorus atoms of the proposed helical model for Structure A are related by diad axes at $z = \frac{1}{4}c$ and $\frac{3}{4}c$, in addition to the diad axes at $z = o$ and $z = c/2$. In Fig. 6 the strength of this psuedo-halving of the c-period of the unit cell may be seen by comparing the sections 0 to 7, of Fig.4 (i), with the sections 15 to 8, of Fig.4. (ii), after the latter have been reflected across a diad axis along b at the level of section 7.5. The Patterson diagram of only that part of the structure that was related by this psuedo-halving was therefore constructed. This was done by superimposing the original Patterson sections 0 to 7 on the reflected sections, 15 to 8 respectively, and tracing the common positive regions. Peaks in the negative regions were also indicated. The eight Patterson sections that resulted from this procedure should contain all of the intra-helical P.P vectors for the proposed model. Therefore the minimum function method was applied to those sections, using displacement vectors having co-ordinates of the theoretical

Fig. 6. (i) The a - b sections from 15 to 8, of the three dimensional Patterson function of Structure A, reflected across a diad axis at the level of section 7.5. The crosses in the figure are defined in Fig. 4 (ii).



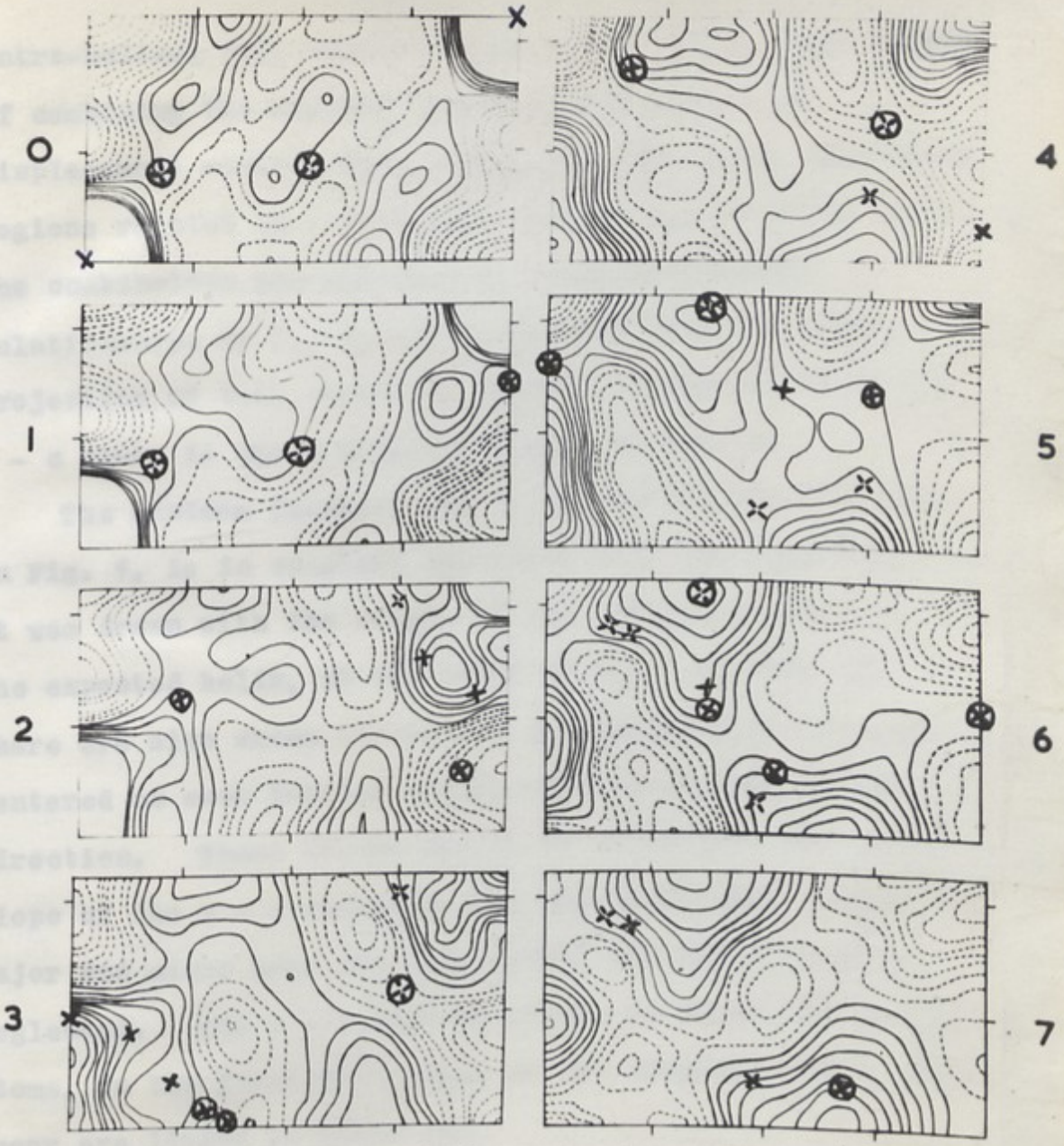


Fig.6. (ii) Comparison of these a - b sections 0 to 7, of the three dimensional Patterson function of Structure A, with the reflected sections 15 to 8 respectively, (shown in Fig.6. (i)) shows the tendency towards a halving of the c period of the unit cell. The crosses in the figure are defined in Fig.4. (i).

intra-helical P.P. vectors shown in Table II. The result of combining the minimum function maps due to the displacement vectors P_oP_3, P_oP_4, P_3P_4 and using only those regions related by a diad axis through P_o, is shown in Fig.7. The combination was effected by using the spacial relationships of P_o, P_3 and P_4 deduced from the model. The projection of this suggested helical configuration on the a - c plane is shown diagramatically in Fig. 8.

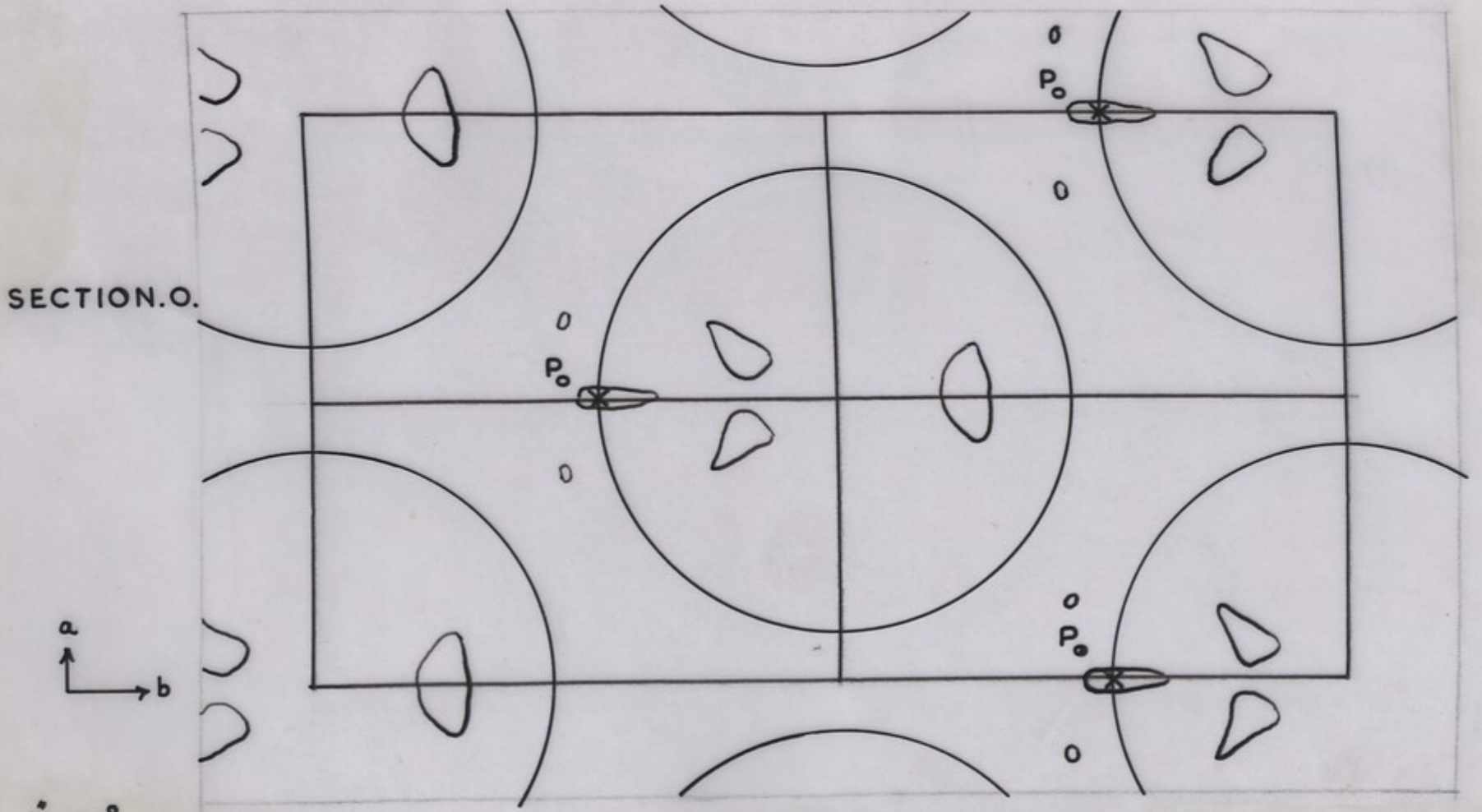
The minimum function (M_6 on Buerger's notation) shown in Fig. 7. is in complete agreement with the suggested model. It was drawn with the origin of the lattice on the axis of the expected helix, at the level of P_o. On each section there are also shown the traces of cylinders of 9A° radius, centered on each lattice point, with their axes in the c direction. These traces should be elliptical due to the slope of the a - b sections, but the difference between the major and minor axes was only 0.06A° and was therefore neglected. The a - b section levels on which the phosphorus atoms, in the first 8A° length of the proposed model, should occur are listed in Table III.

Table III.

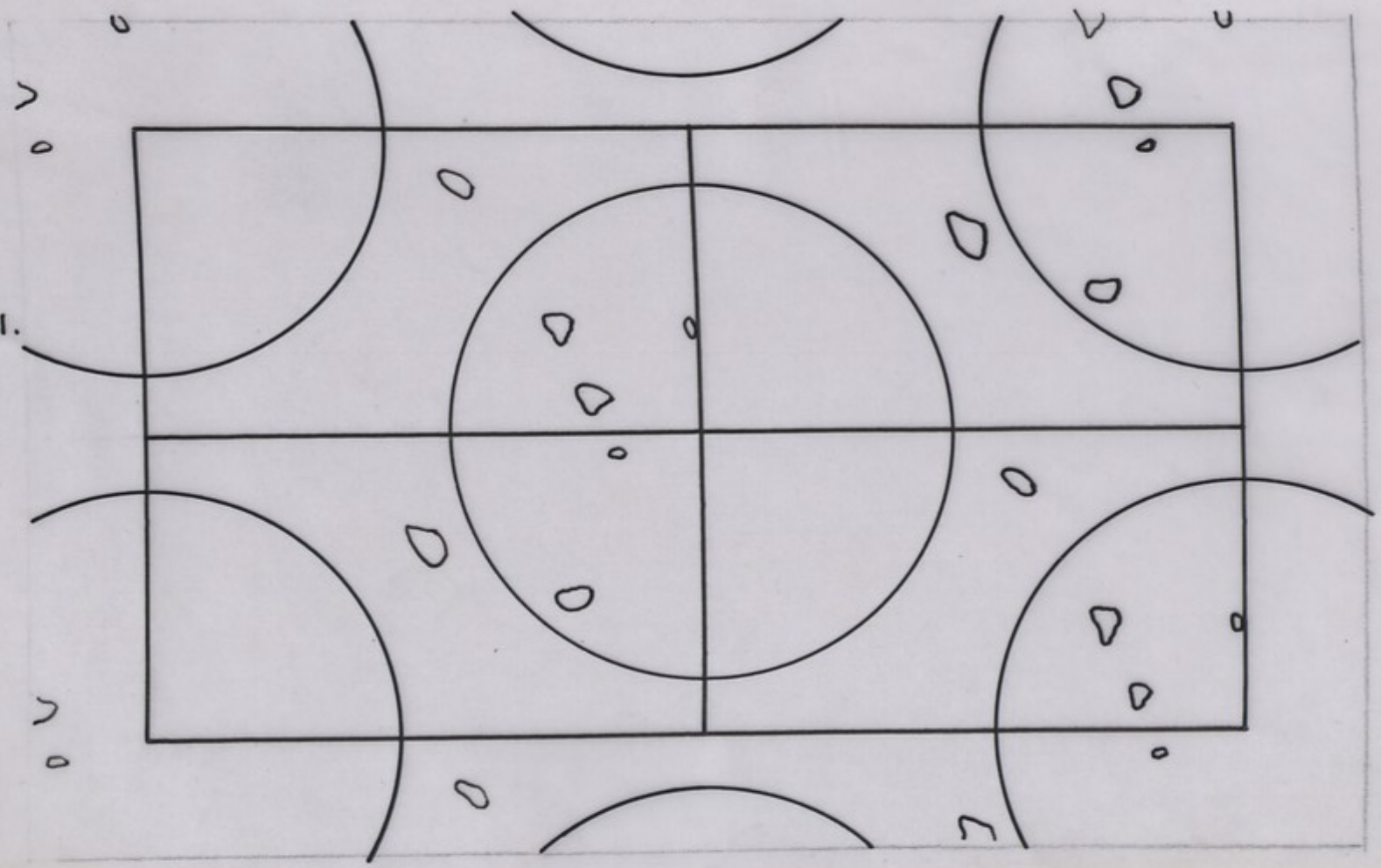
P_n or $P_{n'}$	P_o	P_6'	P_1	P_2	P_7'	P_3	P_8'
Section	0	1.7	2.1	4.5	4.9	7.1	7.9

The expected positions of these phosphorus atoms are marked on the appropriate sections with a full cross except for P_2, the co-ordinates of this atom are indicated on both sections 4 and 5 by a broken cross. As may be seen from

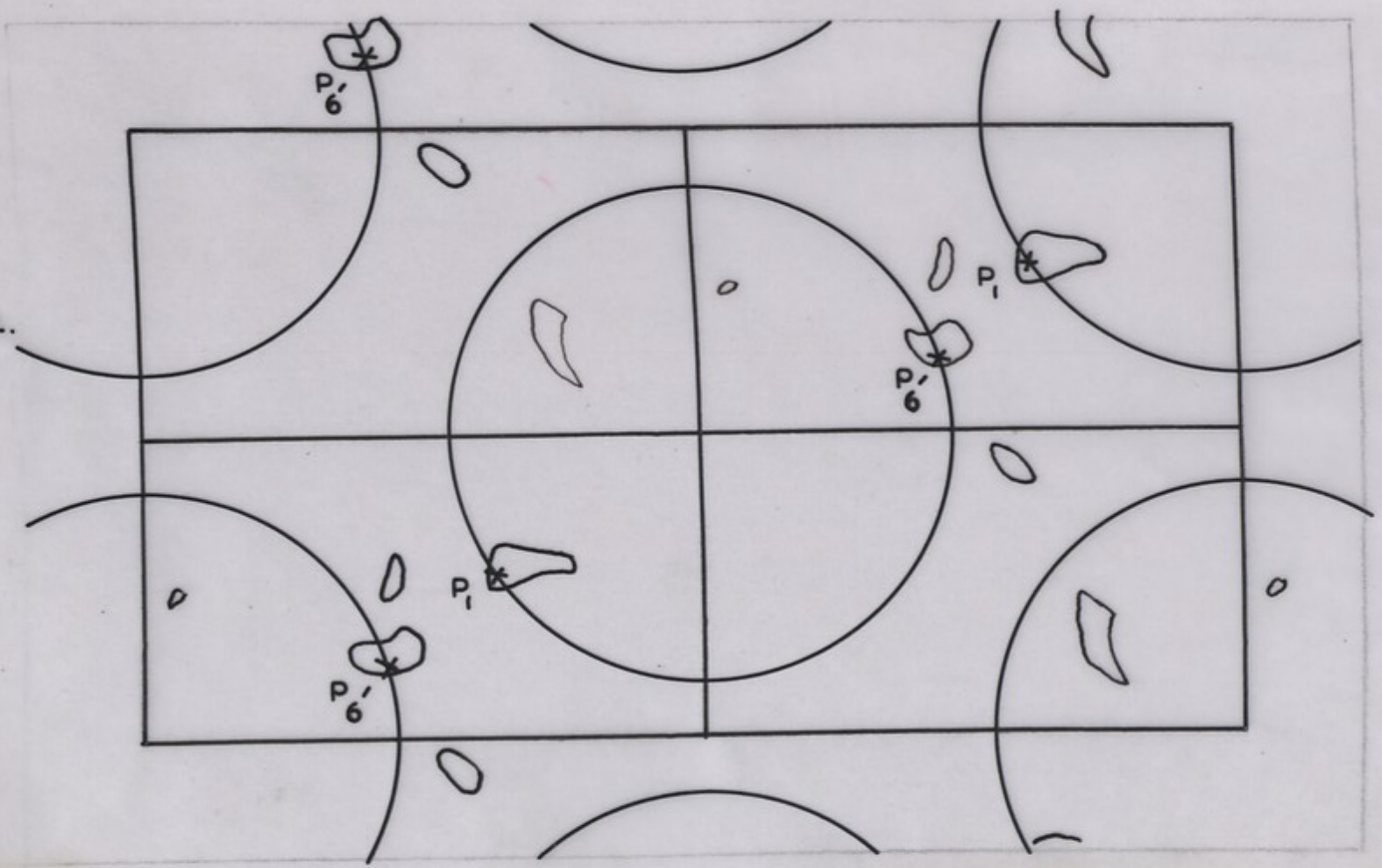
Fig. 7. A M_6 minimum function map of that part of Structure A which repeats in the period $c/2$ and is related by a diad axis in the b direction at $c/4$. (i.e. Section level 7.5).


SECTION.O.
O
Po
O
O
Po
O
O
Po
O
a
b
·1" ≡ ·6A°

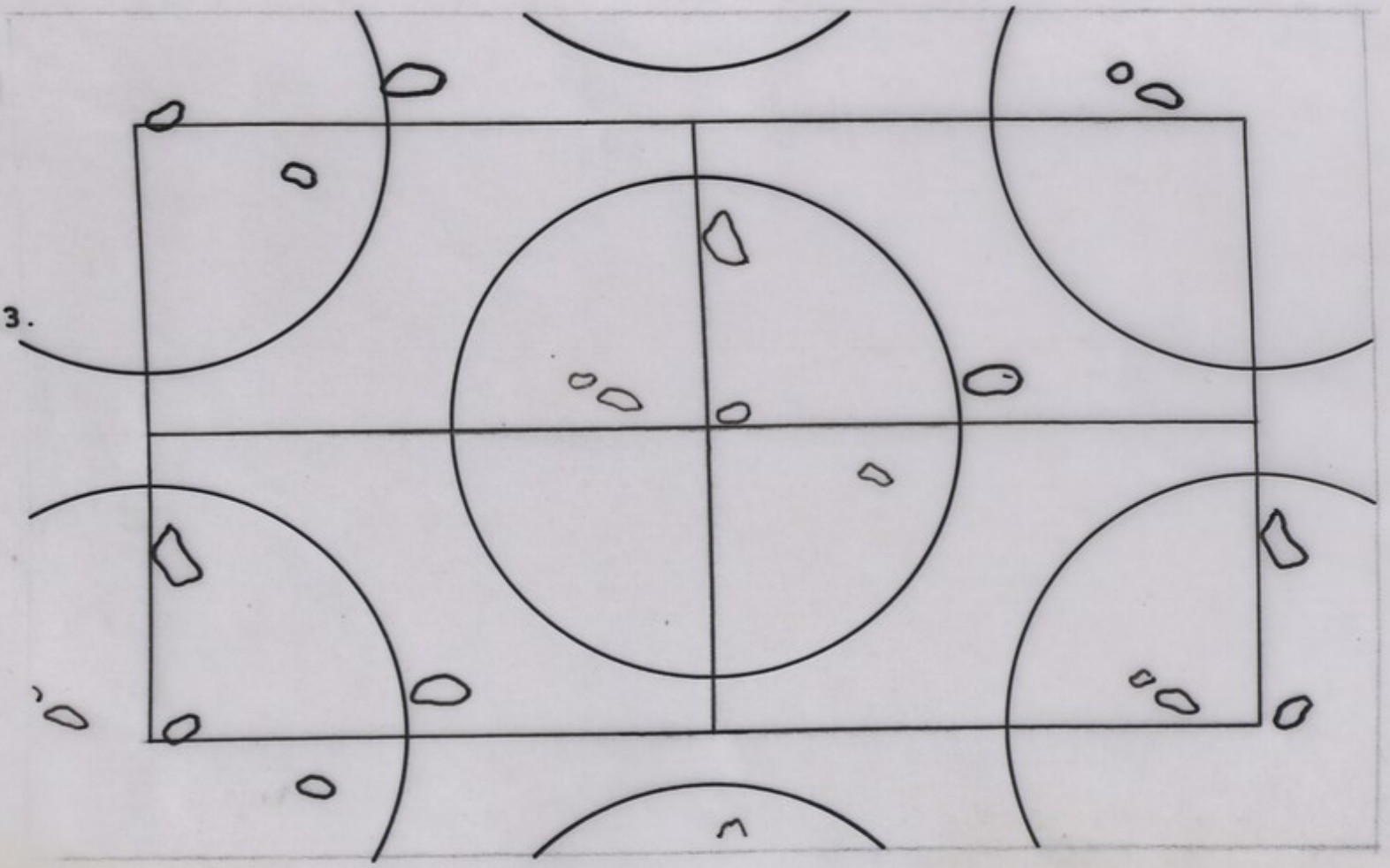
SECTION. I.



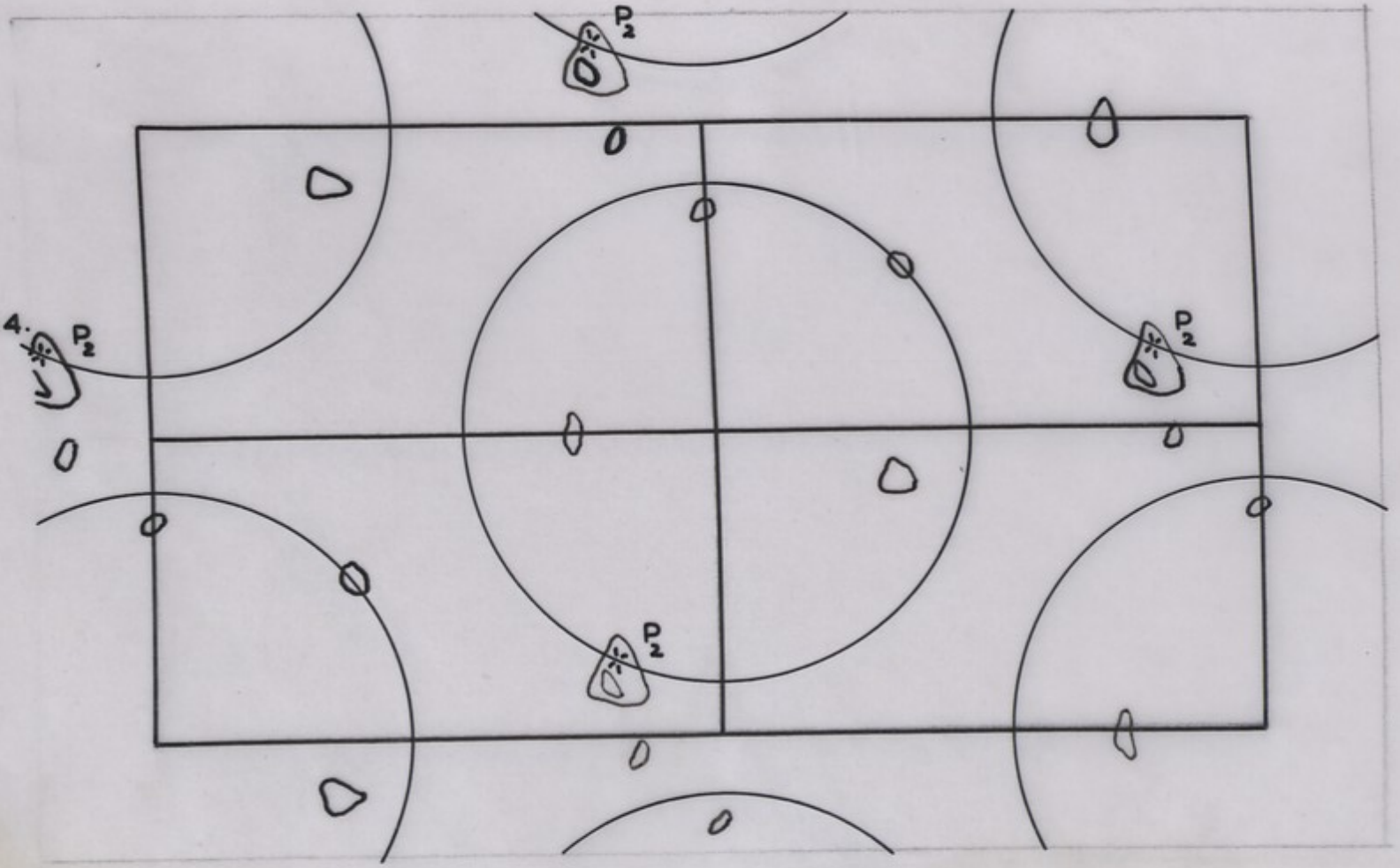
SECTION. 2.



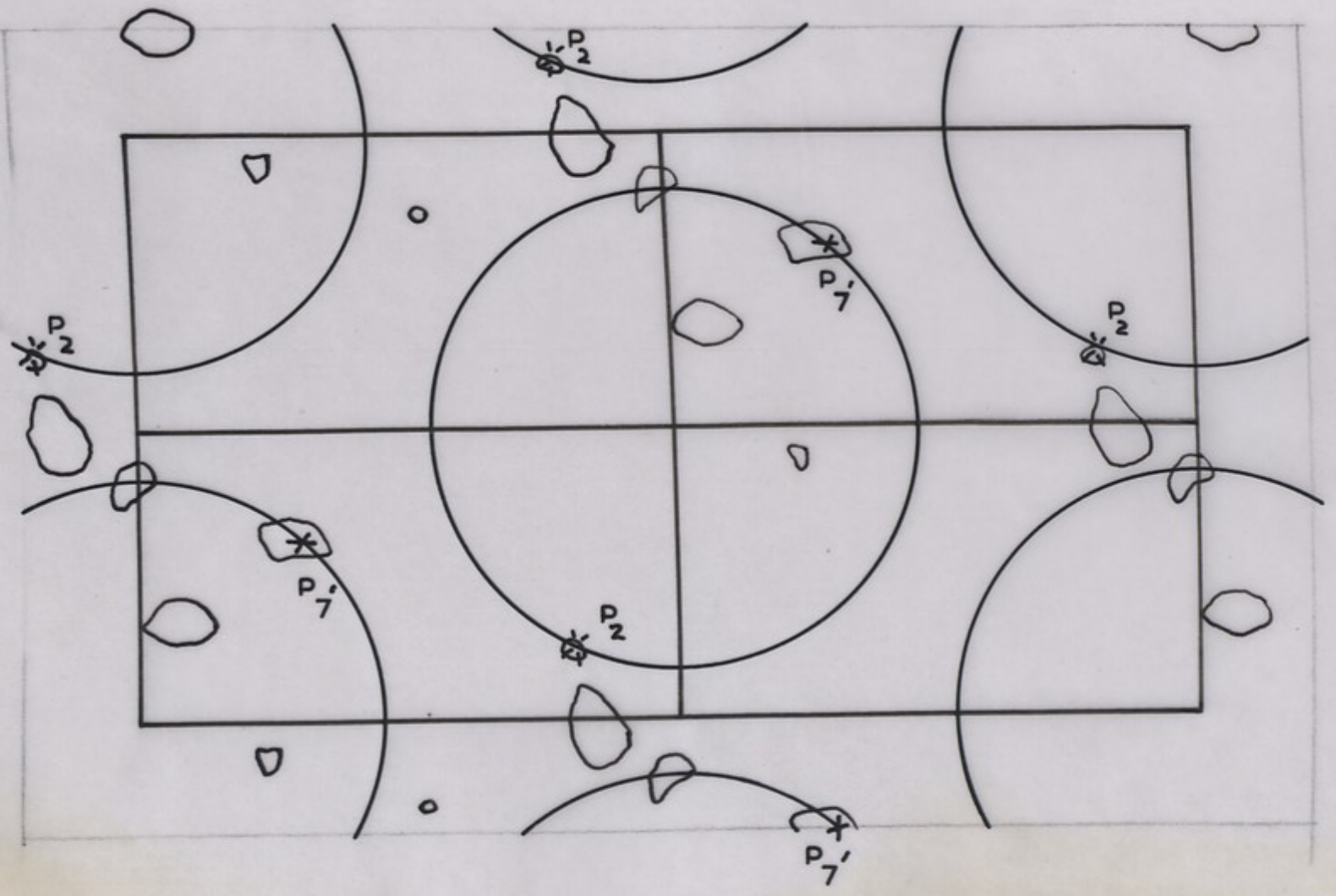
SECTION. 3.



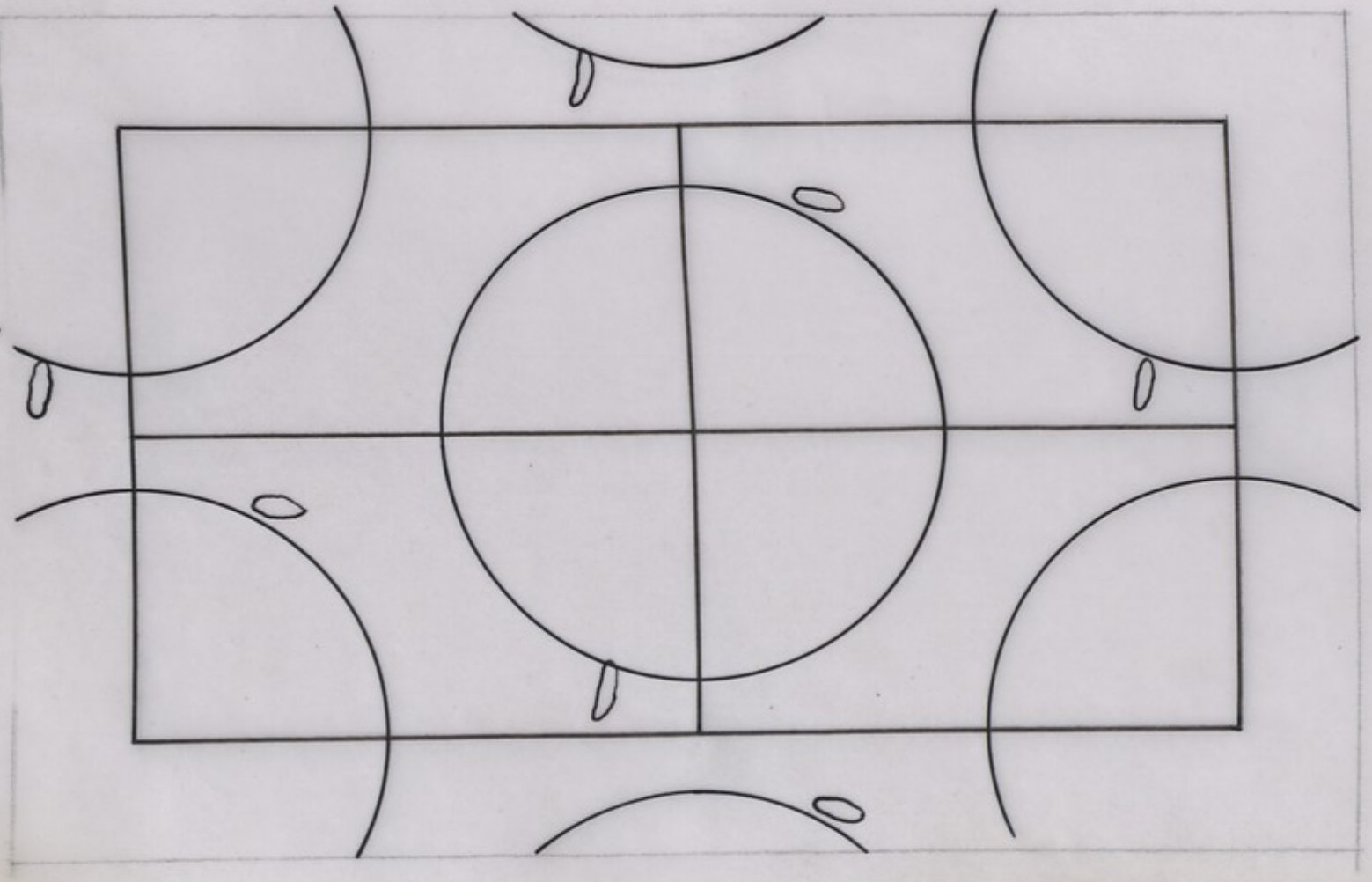
SECTION. 4.



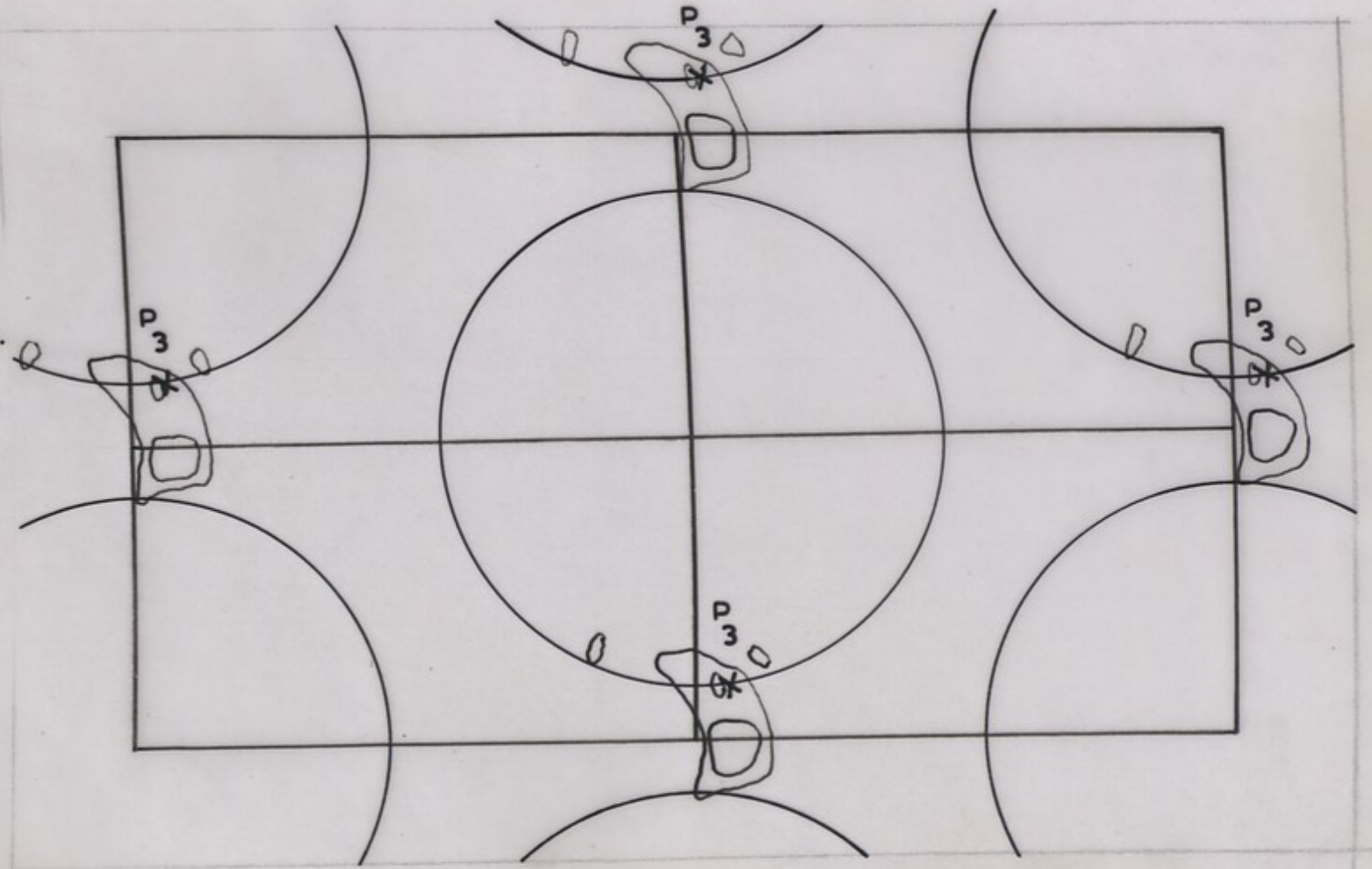
SECTION 5



SECTION . 6 .



SECTION . 7 .



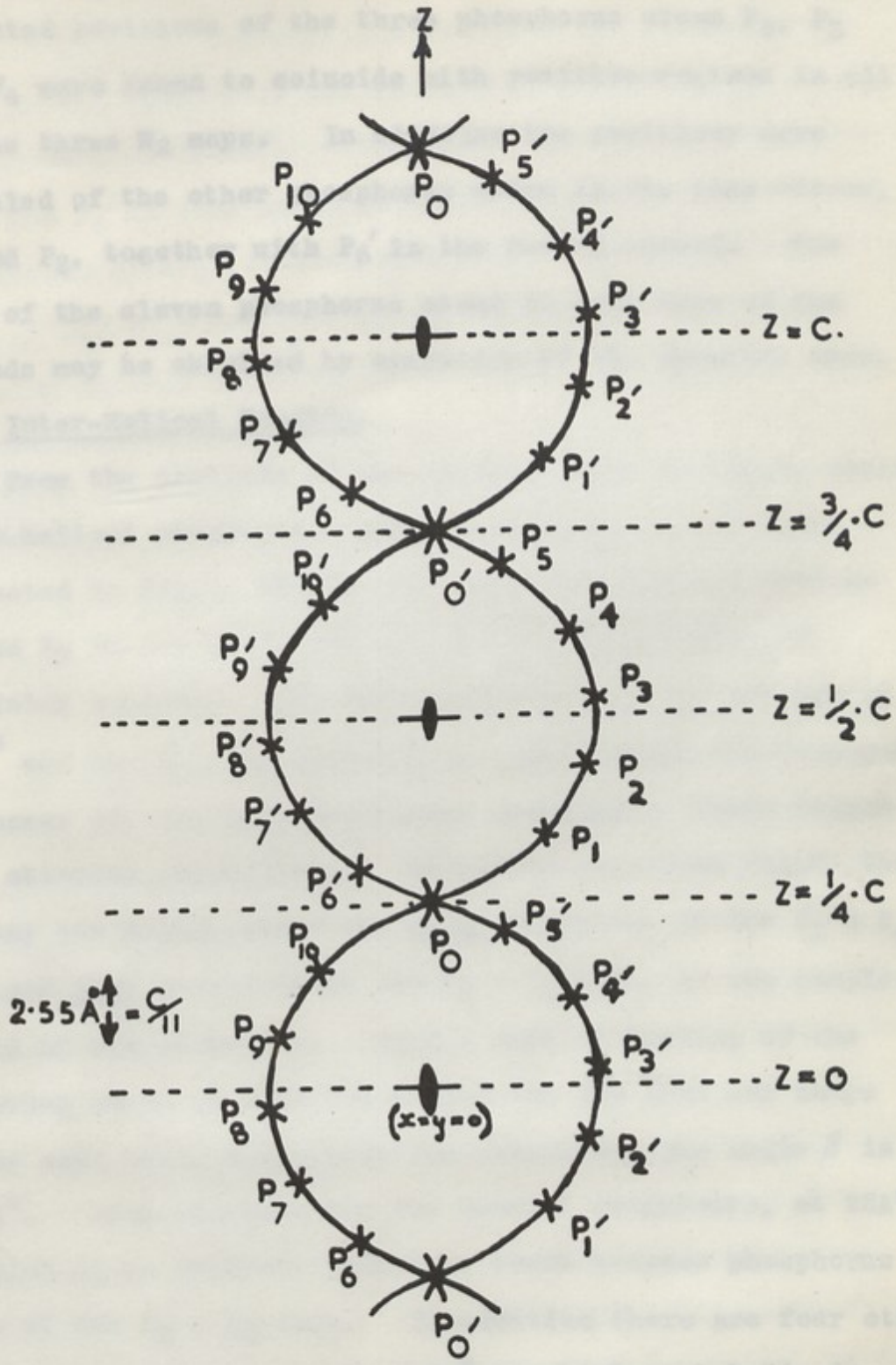


FIG 8: A SCHEMATIC PROJECTION, ON THE A-C PLANE, OF THE CONFIGURATION OF PHOSPHORUS ATOMS FOUND FOR STRUCTURE A. IN ONE HELICAL UNIT.

Fig.8, the phosphorus atom P_4 is equivalent to P_7', thus the expected positions of the three phosphorus atoms P_0, P_3 and P_4 were found to coincide with positive regions in all of the three M_2 maps. In addition the positions were revealed of the other phosphorus atoms in the same strand, P_1 and P_2, together with P_6' in the second strand. The rest of the eleven phosphorus atoms in each turn of the strands may be obtained by operation of the symmetry axes.

(iv) Inter-Helical Bonding.

From the sections of the helices shown in Fig.7. strong inter-helical phosphate - phosphate bonds, of the type suggested in Fig.1. Chapter IV, might be expected between P_1 and P_3 in one helix and P_6' and P_8' respectively, in adjoining helices. The inter-helical $P_1 - P_6'$ distance is 4.9A° and the $P_3 - P_8'$ distance is 4.4A°, which are reasonable distances for the type of linkage suggested; these values were obtained graphically. It may be seen from Fig.8. that, for any one helix, there are eight positions of the $P_1 - P_6'$ type and four positions of the $P_3 - P_8'$ type, in one complete period of the structure. Such a mode of linking of the molecules shows clearly the reason for the size and shape of the unit cell, especially the reason why the angle β is 96.5A°. Each molecule has two nearest neighbours, at 22A°, to which it is probably linked by bonds between phosphorus atoms of the $P_3 - P_8'$ type. In addition there are four other neighbouring molecules, at 22.7A°, to which each helix is probably bonded by linkage between atoms of the $P_1 - P_6'$ type.

These four helices surrounding the helix in the face-centered position have "a" translations relative to that helix of 11A°, which, because of the value of β, produce a z displacement of 1.25A°. The difference in z co-ordinates of P_1 and P_6' in each helix is 1.27A°, therefore this translation of a/2, coupled with the value of β, ensures that P_1 (P_{10}', P_1', P_{10}) and P_6'(P_5, P_6, P_5') in the appropriate adjoining helices are brought to about the same level in z. The difference in the z co-ordinates of P_3 (P_3') and P_8' (P_8) in adjoining helices at 22A°, is 1.2A°, but the number of these possible inter-helical linkages per turn of the helix is only half that of the P_1 - P_6' type.

Inspection of sections 4 and 5 of Fig. 7. suggests that the phosphorus atom P_2 is also engaged in bonding to the same adjoining helix as P_3. The nearest inter-helical phosphorus atom to P_2 is P_8' at a distance of 7.8A°, a phosphate-phosphate linkage between these groups would therefore seem unlikely. The positive region to which P_2 appears to be linked is on the circumference of the trace of the nearest helix at the level of section 5. (difference of about 1.1A° in z) and is distant about 5.5A° from P_2. Now the work of Cavalieri[5.] on the spectrophotometric titration of DNA with acids and magnesium ions has indicated that some of the bases may be bonded to phosphate groups. Brown and Seeds[6.] have suggested that the 2-amino group of guanine can approach near enough to a phosphate group, in the same helix, to form a hydrogen bond of the type P=O .... H - N. It may be,

therefore, that there is hydrogen bonding between a P_2 (P_9', P_2', P_9) phosphate - sodium - water complex and a guanine base in an adjoining helix. This would necessitate a guanine base four times in each complete turn of a helix, i.e. one in 5.5 of the bases must be guanine. This quantitative prediction is in accord with the results of chemical analyses, which show that the mole fraction of guanine to total bases is 0.21 in calf-thymus DNA. Furthermore, the work of Cavalieri has shown that about 1 mole of magnesium ions per 5 moles of phosphorus is bound to calf-thymus DNA and that magnesium probably forms a chelated complex between an amino group and a phosphate group. As Brown and Watson[7.] have pointed out these data suggest that magnesium and similar ions can form a chelate bridge between the 2-amino group of guarine and a neighbouring phosphate group. These authors[7.] advanced the suggestion of linkage between the 2-amino group of guanine and a phosphate group in the backbone chain of the same helix as an explanation of the differences in binding properties found for different fractions of calf-thymus DNA with histone. It has been suggested, in Chapter IV, that the Structure B helical form probably persists in solution, so that the above suggestion of Brown and Watson would apply to the B helix. When the structure contracts to the A form, it may be that the suggested intra-helical base-phosphate linkage is broken and cannot reform due to the inter-helical phosphate-phosphate bonding, so that a base-phosphate bridge between neighbouring

helices is formed. Such additional inter-helical bonding, in Structure A, close to the "a" direction might explain why the separation of the helices is least in that direction.

REFERENCES IN CHAPTER VII

1. Watson, J.D. and Crick, F.H.C. (1953) Private Communication.
2. Buerger, M.J. (1950) Acta Cryst. 3. 87.
3. Buerger, M.J. (1951) Acta Cryst. 4.
4. For references and a critical discussion of these methods see Buerger.[3.]
5. Cavalieri, L. (1952) J. Amer. Chem. Soc. 74. 1242.
6. Brown, G.L. and Seeds, W.E. (Unpublished).
7. Brown, G.L. and Watson, M. (1953) 172. 339. Nature.

SUMMARY OF THE EVIDENCE PRESENTED, IN CHAPTERS VI AND VII, ON THE MOLECULAR CONFIGURATION IN STRUCTURE A

From the study of the cylindrical Patterson function, presented in Chapter VI, the model outlined below has been put forward as the configuration of a molecule of NaDNA held in the crystalline state in fibres of that material.
Proposed Structure: The phosphorus atoms from two polynucleotide chains form two co-axial helical strands of 9A° radius and 28.1A° pitch, with 11 phosphorus atoms spaced equally along each turn of each strand. The separation of these helical strands in the direction of their common axis is 14.05A°. Corresponding successive phosphorus atoms in each strand, P_n and $P_{n'}$, have the same co-ordinates in a plane at right angles to the axis of the helix, which is coincident with the c direction in the crystallites. (fibre axis direction).

Density measurements and the unit cell dimensions have shown that the sugar and base residues, of the proposed two polynucleotide chain helix, must be turned inwards toward the axis of the helix. In addition, the space group found for the unit cell contents requires that the sequence of the atoms in the two phosphate-sugar chains should run in opposite directions, with respect to the axis of the helix.

The phosphorus-phosphorus vectors to be expected from the proposed model have been found to be in agreement with the three dimensional Patterson function of the crystallite structure. Application of the minimum function method of Buerger, discussed in Chapter VII, to the three dimensional

Patterson function of only that part of the crystallite structure which repeated in the period $c/2$, led to the construction of an M_6 minimum function map which was found to be in complete accord with the proposed structure.

If it may be considered that the important inter-helical linkages are phosphate-phosphate bonds, then it has been shown (Chapter VII) that the proposed structure readily accounts for the size and shape of the unit cell. The M_6 minimum function map, mentioned above, has indicated bonding between nearest phosphorus atoms in adjoining helices and has indicated also an additional inter-helical bonding which may be interpreted as a 2-amino-guanine --- phosphate linkage.

CHAPTER VIII

THE CORRELATION OF THE HELICAL MOLECULE SUGGESTED FROM THE X-RAY WORK, WITH OTHER PHYSICAL STUDIES OF DNA.

Increasing the water content of NaDNA fibres beyond the lower limit necessary for the A to B transformation increases the lateral swelling of the fibres. There is no further appreciable length change once the A to B transition has occurred, furthermore, there is no change in the general form of the X-ray diagram. It would seem, therefore, that Structure B may quite possibly be close to the molecular configuration of NaDNA in solution. If this is true, then a definite model is available on which to base the interpretation of the results of other physical studies of DNA.

Atomic Models.

The first detailed atomic model for DNA was that proposed by Pauling and Corey.[1] The X-ray diagrams on which the model was based were fibre diagrams similar to those of Astbury, which have been shown here to be a mixture of the two phases A and B. The suggested structure was a three-strand helix in which the phosphate groups formed a dense core. It is difficult to see how the swelling and solution of DNA in water could be explained by such an arrangement. Moreover, if such a structure existed in the crystalline state it would require to undergo a radical intra-molecular re-arrangement in passing to the wet state. For even if the interpretations of the two types of X-ray diagram presented here are discounted, there still remains to explain the accessibility of the phosphate groups and the inaccessibility of the NH_2 and CO groups during titration. In addition,

the phosphate groups must be in an accessible position for interaction with proteins. Therefore, since the transition from the crystalline to the wet state is both readily and rapidly reversible, the structure proposed by Pauling and Corey for the crystalline state would seem unlikely.

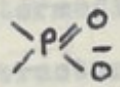
The only other detailed atomic model for DNA is that which has been put forward by Watson and Crick.[2] From the qualitative discussion in Chapter IV and the measurements discussed in Chapter V, the molecular configuration of Structure B appears quite consistent with the general features of this model. On matters of detail there is disagreement. The radius of the phosphorus "helix" would appear from the X-ray data to be closer to 8.5A° than to the 10A° suggested in the Watson and Crick model. Also the helices are probably slightly off circular in cross-section.

The X-ray data have not sufficient resolving power to say anything about the nature of the inter-base arrangements. The X-ray diagrams, of both Structure B and Structure A, can be convincingly interpreted on the basis of particles consisting of a helical array of the heavy phosphate groups from two polynucleotide chains, with their sugar and base groups turned inwards toward the helical axis. That it is the hydrogen bonding of the base groups that links these two helically arranged polynucleotide chains together, can only be argued from titration work such as that of Gulland and Jordan.[3] With regard to the base sequence in the structure, the crystallinity of Structure A would seem to suggest

that the bases were not randomly arranged, i.e. equivalent crystallographic positions were not occupied by purines and pyrimidines. This would imply a restriction to the biological specificity. The Watson and Crick model overcomes this difficulty by postulating that the bases are always combined as purine - pyrimidine pairs, the base-sugar linkages to the two helical phosphate-sugar chains being related by a diad axis lying in the plane of this base complex. Therefore a crystallographic array of such complexes would still permit the individual base sequence to vary widely.

Structure B in Solution - Line density and Bound Water.

In order to calculate the shape and size of particles in Structure B form to be expected from any experiments on solutions of NaDNA, it is first necessary to decide on the amount of water that is likely to be bound by this structure. The major part of any bound water is most likely to be associated with the phosphate groups around the outside of the molecule, and so would be expected to vary with the degree of dissociation of these groups. Therefore the effective line density of the NaDNA structure will probably vary with the pH and the ionic strength of the solution. Consider the possible values for the undissociated structure. The average molecular weight of a sodium nucleotide is 330, therefore $6.6.10^3$ is the molecular weight of a 33A° length of an anhydrous Structure B particle, i.e. a line density of 200. (Molecular weight/A°). If the bound water were 43.6% of the dry weight, i.e. 8 molecules of water per nucleotide, the line density would be 287 M/A°.

The work of El-Sabeh and Hasted[4] has shown that in neutral aqueous solution, if the sodium phosphate groups were completely dissociated, the amount of water found by dielectric measurements to be irrotationally bound to the NaDNA particles was only of the order of 10% of the dry weight. This corresponded to a binding of just less than two water molecules per nucleotide. However, Hasted has pointed out that if we consider the phosphate groups in the NaDNA structure to be in the dissociated form we are dealing with the group  , work on the binding of water by Sulphate ions[5] would then lead us to expect about two irrotationally bound water molecules per single bonded oxygen, plus the rotational binding of a further two water molecules. This suggests that the phosphate group of each dissociated nucleotide may bind four molecules of water, which would give a completely dissociated Structure B form a line density of 230 M/A°. This figure, however, depends on the assumption that, under the conditions of the experiment to measure the dielectric properties of the NaDNA solution, all the sodium phosphate groups were in fact dissociated.

Interpretation of Light Scattering Data in Terms of Structure B.

The particle dimensions calculated from the results of light scattering depend on the initial assumptions made as to particle shape. If the above suggestion is correct i.e. that "Structure B" is the molecular configuration to be expected of NaDNA particles in solution, a definite model now exists on which to base such calculations. In what follows

the light scattering data will therefore be examined with a view to interpretation on the basis of Structure B, in conjunction with the detail of the Watson and Crick atomic model.[2]

The light scattering observed depends partly on the difference in density of the scattering material and the solvent. Therefore in any considerations of the shape and size of particles in Structure B form, based on light scattering data, the line density should probably be taken as something intermediate between the value of 200 M/A°, for anhydrous Structure B, and the value of 287 M/A°, which corresponds to the binding of 8 molecules of water per sodium nucleotide.

The length to be expected from a Structure B molecule having a molecular weight of $7.3.10^6$ and a hydrated line density of 250 M/A° is 29,200A°. Therefore from Steiner's[6] results, (Table 1), for the length of particles found assuming "stiff rods", it would appear that single Structure B particles do not exist in solution in an axially straight form. However, side by side aggregation of four axially straight Structure B particles having a length of 7,300A° would account for the molecular weight, and would be in reasonable agreement with the experimental value of 7,180A° found by assuming monodispersed rods. Such a grouping would be essentially a rod of length 7,300A° and diameter about 50A°. This may be considered, in relation to the wavelength of the light used in these experiments, as a rod of negligible thickness and the dissymmetry factor for such

TABLE I

Results from Light Scattering Measurements on two samples of calf-thymus NaDNA in 0.2M. KCl. by R.F. Steiner. 6.

Molecular Weight (15%).	Particle Dimension (10%).	
$4.4.10^6$	L_1	4730A°
	L_2	3880A°
	L_3	6710A°
$7.3.10^6$	L_1	5080A°
	L_2	4160A°
	L_3	7180A°

L_1 = root mean square end to end separation calculated on basis of monodisperse randomly kinked coils.

L_2 = root mean square end to end separation calculated on basis of polydisperse random coils.

L_3 = length calculated on basis of monodisperse rods.

7.8.9.

a rod is 2.46. Recent reports on light scattering all give values greater than 3.0 for the observed dissymmetry factors of the scattering particles in several different preparations. Therefore the high molecular weights and the scattering curves observed cannot be explained in terms of axially straight Structure B particles.

We must then, consider the possibility of the helix of Structure B existing in some macro-coiled state. The various reports quoted above seem to agree that the observed data fit quite closely to a system of polydispersed, rather than monodispersed, randomly kinked coils. Assuming such a system, for a preparation of molecular weight of $7.3.10^6$ in neutral solution, the R.M.S. end to end separation of the coils has been found by Steiner[15] to be 4160A°. (see also results of Doty et al[10] Table II) which in terms of Structure B means that a very long molecule, (29,000A°) has achieved an end to end separation of about one-seventh of its total length. This indicates a degree of folding necessitating a rather flexible structure. However, the asymmetry of the Structure B molecule is so high that if this end to end separation were due to the molecule forming an arc of a circle, the deviation between the normals to the cross-sections at each end of one turn of the helix would be negligible e.g. for length 29,000A° and end to end separation of only 2,416A°, the radius of curvature of the circle is as large as 5,000A°. In addition to the "asymmetry effect"

the flexibility could be further increased by breaking the cross-linking hydrogen bonding between the base rings, at various points along the helix. Once this were done, the free rotation about a number of bonds in the inter-nucleotide links of the phosphate-sugar chain would permit further macro-coiling.

Gulland and Jordan[3] have shown that outside the range pH 5 - 11 the CO and NH_2 groups of the bases become titratable. Therefore if the rupture of hydrogen bonding between these groups was playing an important part in determining the flexibility of the molecule, it would be expected that over the range pH 11 - 5 there would be no appreciable change in the R value. (i.e. end to end separation of the particle). Outside that range the R value should decrease progressively with pH due to an increased folding of the molecule as more and more hydrogen bonds were broken. The molecular weight should stay constant until the pH became sufficiently low for primary bonds in the phosphate-sugar backbone to be broken, so causing depolymerisation. The recent work of Doty et al[10] strongly supports these suggestions. Measurements on NaDNA solutions at pH 6.5, 3.0 and 2.6 were reported. It was found that the R value decreased on lowering the pH whilst the molecular weight stayed constant, (Table II), and that the process was reversible. Doty pointed out that since the solutions contained 0.2M NaCl, less amino-groups

TABLE II

Results of Doty et al[10] from Light Scattering measurements on Sodium Thymonucleate in 0.2M NaCl showing the changes induced by dilute acid.

The following data are from the same specimen; the molecular weight remained constant, throughout the pH variation, at a value of $7.7.10^6$.

pH.	R_A^0 (Å)
6.5	4110
3.0	3550
2.6	1760
6.5 (return)	4250

R = Root mean square end to end separation of the particles calculated for polydisperse unbranched random coils. The polydispersity corresponding to a ratio of weight to number average molecular weight of two.

N.B.
The NaDNA sample was prepared by following the method of Signer and Schwander.

would become charged at the lower pH than might be expected from titration data in the absence of salt. It follows that most of the inter-base hydrogen bonds of Structure B would still be intact. Therefore a complete breakdown of the helix would not occur, but sufficient hydrogen bonds would be broken to cause an appreciable increase in the flexibility of the molecule. The fact that the results were reversible with pH, indicates that these inter-base linkages can reform easily.

The more quantitative approach of Peterlin,[11] outlined in Chapter III, for a Signer preparation of calf-thymus sodium nucleate having a molecular weight of $6.7.10^6$, has lead to a line density of 156 M/A°. This value was derived by matching an experimental scattering curve with one of a family of theoretical curves for the scattering to be expected from a continuous thread-like flexible particle at various values of the flexibility, there being negligible polydispersity for any one value. This latter assumption may not be true, which may possibly account for the difference between the value of the line density derived by this method and that to be expected from a hydrated "Structure B".

Ultracentrifuge Data.

(In what follows the diameter of a hydrated Structure B particle will be taken as 20A°).

The particle parameters calculated from their experimental

data by Tennent and Vilbrant[12] were; molecular weight $5.8.10^5$, length 5,200A° and shape factor 400. For this molecular weight, even if the particles had the minimum unhydrated line density for Structure B of 200, the length would only be 2,900A° and the shape factor 145. Some of the difference in these values of the shape factor may possible be due to computing the experimental values from the Perrin equation[13] which, as Tennent and Vilbrant remark, probably does not hold for molecules as asymmetric as NaDNA.

In the work reported by Cecil and Ogston[14] the values given for the shape factors are too low. These authors found a shape factor of 140 for a molecular weight of 10^6, and 170 for $1.3\ 10^6$ molecular weight. For single particles to have these values the hydrated Structure B line density would require to be 357 M/A°, and 382 M/A° respectively, if the systems were monodisperse. It would seem, therefore, that particle aggregation might have occurred.

Apart from any aggregation effects which might occur when these long particles are orientated by the high centrifugal fields used in these experiments, it may be that the molecules themselves are extended.

Helical Structure Extensible.

That the structure of NaDNA is extensible has been suggested by the observations of Wilkins[15] on the macroscopic behaviour of NaDNA fibres. It has been shown that negatively birefringent fibres, whose X-ray pattern may be either A or

B, can be stretched with an increase of birefringence until sudden necking occurs. The author has found that the force required for this is very small, less than 0.1 gm. for a 40μ diameter fibre. Further extension of the fibre takes place by the two shoulders of the necked region separating and travelling along the fibre, the extended region between the shoulders exhibiting very small positive birefringence. This result, together with the fact that the process may be reversed if the fibre is allowed to take up water, (which it does very rapidly), suggests that the helical molecule is capable of being extended, the base rings tilting towards the axis of the helix.

The sign change of the birefringence, from negative to positive, may be caused in the sheet specimens by rolling them. X-ray diffraction patterns, obtained by the author, of such positively birefringent sheet specimens showed no evidence of regular structure. This is consistent with a rather variable extension of the individual molecules in the specimen.

Flow Birefringence.

Edsall[16] has applied an approximation of the Perrin equation to the early work of Signer et al[17] and that of Kausche et al[18]; assuming a shape factor of 200 the approximate length of the particles was given as 4,500A°. The associated

molecular weight was said to lie between ½ and 1 Million. If the molecular weight were 1.10^6 then the line density of a particle 4,500A° long would be 220 M/A°. This value is very close to that expected of a hydrated Structure B configuration. However, in the recent work of Schwander et al[19] values were reported of 8000A° and 10A°, for particle length and breadth respectively. If it be accepted that the molecule may be extended, then perhaps this greater assymmetry, found by Schwander et al, might be due to a stretching of the molecule by the tensions in the flowing liquid. Such a process has been found by Signer and Sadron[20] working with polystyrenes of high molecular weight.

Electron Microscopy.

The studies of Williams[21] and of Kahler et al,[22] mentioned in Chapter II, afford a direct confirmation of the molecular diameter of DNA suggested by the interpretation of the X-ray data presented here. It may be significant, with regard to the difference between the molecular diameter of 17A° suggested here for Structure B and the 20A° diameter of the model structure put forward by Watson and Crick, that the diameter of the fibrils observed by the workers mentioned above[21.22.] are in each case smaller than 20A°. ("15A° approximately"[21] -- "in the range 10 to 20A°"[22.]).

Conclusion.

From the above discussion it may be concluded that, in general, where the data permit of the determination of any molecular parameters these are consistent with the existence, in solution, of the NaDNA molecule in the form found in Structure B.

REFERENCES IN CHAPTER VIII.

1. Pauling, L. and Corey, R.B. (1953) Proc. Nat. Acad. Sci. 39. 84.
2. Watson, J.D. and Crick, F.H.C. (1953) Nature. 171. 737.
3. Gulland, J.M. and Jordan, D.O. (1947) Cold Spring Harb. Symp. Quant. Biol. 12. 5.
4. El-Sabeh, S.H.M. and Hasted, J.B. Private Communication.
   El-Sabeh, S.H.M. (1953) Ph.D. Thesis London.
5. Hasted, J.B., Ritson, D.M. and Collie, C.H. (1948). J. Chem. Phys. 16. 1.
6. Steiner, R.F. (1952) U.S.A. N.M.R.I. Memorandum Report. 52-2. 10. 947.
7. Doty, P. and Bunce, B.H. (1952) J. Amer. Chem. Soc. 74. 5029.
8. Rowen, J.W. (1953) Biochimica et Biophysica Acta. 10. 391.
9. Smith, D.B. and Sheffer, H. (1950) Canadian. J. Research. B.28. 96.
10. Reichmann, M.E., Bunce, B.H. and Doty, P. (1953) J. Polymer. Science. 10. 109.
11. Peterlin, A. (1953) Nature 171. 259.
    (1953) J. Polymer Sc. 10. 425.
12. Tennent, H.G., and Vilbrant, C.F. (1943) J. Amer. Chem. Soc. 65. 424.
13. Perrin, F. (1936) J. Phy. Radium. 7. 1.
14. Cecil, R. and Ogston, A.G. (1948). J. Chem. Soc. 1382.
15. Wilkins, M.H.F., Seeds, W.E. and Gosling, R.G. (1951). Nature. 167. 759.

REFERENCES IN CHAPTER VIII. contd.

16. Edsall, J.T. (1942) Adv. in Colloid Sc. 1. 269.

17. Signer, R. Caspersson, T. and Hammarsten, E. (1938). Nature 141. 122.

18. Kausche, G.A. Guggisberg, H. and Wissler, A. (1939). Naturwissenschaften. 27. 303.

19. Schwander, H. and Cerf, R. (1949) Helv. Chim. Acta. 32. 2356.

20. Signer, R. and Sadron, C. (1936) Helv. Chim. Acta. 19. 1324.

21. Williams, R.C. (1952) Biochim et Biophys. Acta. 9. 237.

22. Kahler, H. and Lloyd, B.J. (1953) Biochim et Biophys. Acta. 10. 355.

APPENDIX

Zero Order Bessel Function. Jo(u).

Calculated for u = 40 - 76. from the approximation formula

$$Jo(u) = \frac{\sin u + \cos u}{\sqrt{\pi . u}}$$

at intervals of u = 0.3. Jo(u) is correct to ± 0.0002.

u	Jo(u)	44.6	.1177	49.4	.$\overline{0091}$	54.2	.$\overline{1084}$
40.058	0	44.9	.1180	49.7	.0244	54.5	.$\overline{1030}$
40.4	.$\overline{0421}$	45.2	.1076	50.0	.0561	54.8	.$\overline{0884}$
40.7	.$\overline{0752}$	45.5	.0879	50.3	.0825	55.1	.$\overline{0660}$
41.0	.$\overline{1010}$	45.8	.0604	50.6	.1011	55.4	.$\overline{0379}$
41.3	.$\overline{1176}$	46.1	.0278	50.9	.1106	55.7	.$\overline{0065}$
41.6	.$\overline{1236}$	46.4	.$\overline{0072}$	51.2	.1102	56.0	.0252
41.9	.$\overline{1185}$	46.7	.$\overline{0413}$	51.5	.1001	56.3	.0546
42.2	.$\overline{1032}$	47.0	.$\overline{0715}$	51.8	.0809	56.6	.0797
42.5	.$\overline{0786}$	47.3	.$\overline{0952}$	52.1	.0549	56.9	.0960
42.8	.$\overline{0471}$	47.6	.$\overline{1102}$	52.4	.0237	57.2	.1046
43.1	.$\overline{0118}$	47.9	.$\overline{1152}$	52.7	.$\overline{0088}$	57.5	.1036
43.4	.0244	48.2	.$\overline{1101}$	53.0	.$\overline{0407}$	57.8	.0937
43.7	.0582	48.5	.$\overline{0950}$	53.3	.$\overline{0688}$	58.1	.0753
44.0	.0866	48.8	.$\overline{0720}$	53.6	.$\overline{0906}$	58.4	.0502
44.3	.1070	49.1	.$\overline{0423}$	53.9	.$\overline{1040}$	58.7	.0210

u	Jo(u)	63.8	.0982	68.9	.0521	74.0	$.\overline{0532}$
59.0	$.\overline{0108}$	64.1	.0882	69.2	.0734	74.3	$.\overline{0369}$
59.3	$.\overline{0401}$	64.4	.0704	69.5	.0882	74.6	$.\overline{0075}$
59.6	$.\overline{0664}$	64.7	.0463	69.8	.0951	74.9	.0263
59.9	$.\overline{0866}$	65.0	.0183	70.1	.0934	75.2	.0511
60.2	$.\overline{0997}$	65.3	$.\overline{0112}$	70.4	.0834	75.5	.0713
60.5	$.\overline{1025}$	65.6	$.\overline{0396}$	70.7	.0660	75.8	.0851
60.8	$.\overline{0968}$	65.9	$.\overline{0644}$	71.0	.0428	76.1	.0912
61.1	$.\overline{0826}$	66.2	$.\overline{0832}$	71.3	.0159		
61.4	$.\overline{0611}$	66.5	$.\overline{0946}$	71.6	$.\overline{0123}$		
61.7	$.\overline{0343}$	66.8	$.\overline{0975}$	71.9	$.\overline{0393}$		
62.0	$.\overline{0045}$	67.1	$.\overline{0917}$	72.2	$.\overline{0626}$		
62.3	.0256	67.4	$.\overline{0778}$	72.5	$.\overline{0804}$		
62.2	.0532	67.7	$.\overline{0569}$	72.8	$.\overline{0907}$		
62.9	.0760	68.0	$.\overline{0311}$	73.1	$.\overline{0932}$		
63.2	.0917	68.3	$.\overline{0026}$	73.4	$.\overline{0871}$		
63.5	.0994	68.6	.0260	73.7	$.\overline{0734}$		